University of Ljubljana
Faculty of Social Sciences
Centre for Social Informatics

Post-stratification weighting of the ESS

Methodology development and data preparation

DACE project, activity of WP12

Vasja Vehovar

Ana Slavec

Marcel Kralj

September 2012

(last update April 29 2014)

Abstract

One of the tasks (Task 3 in WP12) of the DACE project is the elaboration of population weighting to correct for discrepancies with control data. The goal here is to establish methodology and standardized procedures, which will produce data in publicly available format together with corresponding documentation. The task is complex, as we already have data for all countries for all five rounds, with each survey having (some or potential) methodological specifics and problems with control variables. The first deliverable for this task is October 2012 (weights for 5th round), while the second deliverable is April 2013 (all weights R1-R4 and all corresponding documentation).

A basic methodological outline has already been defined in previous weighting attempts, which were temporarily suspended in 2009 because of the problems with ESS data quality, predominantly related do to education (Vehovar et al 2008). Nevertheless, the approach outlined here follows the previous weighting attempts using age, gender, education and region. Similarly, Eurostat LFS data serve as a benchmark for the population variables. However, even if these two most crucial decisions are not changed, numerous open issues arise with the practical production of the weights.

We present here the overview of the current status and we outline the remaining work, particularly data quality process, timing, inventory of problems and outline of reports. We estimate that work is progressing according to the plans. However, there is a set of open methodological questions. Besides numerous country and wave specific problems the main methodical issues are related to the following:

- Confirmation of the existing strategy and practice related to the education variable:
    - Role of ESS education expert interventions into national ISCED,
    - The specifics of the categorizations into low, medium, high,
    - The number of educational categories (low, medium, high),
- The approach to region: using NUTS2, NUTS3 and national specific regions,
- The approach for handling missing data at sample and at population control level,
- The approach to weighting and trimming.

We fully elaborate the above issues here and we also propose solutions, which is to be verified and approved.

Index

1 Introduction and description of work

Given the five ESS rounds, each including from 20 to 30 countries, the task here is complex for two reasons:

- First, there are simply a lot of surveys, more than 100, which all need to be weighted with internal country-consistency (i.e. different rounds for same country) and also with internal round-consistency (i.e. consistent comparisons of different countries for certain round).
- The external consistency also needs to be maintained, i.e. matching with other national surveys and population data sources, national and international. A second source of complexity arises from specifics of each survey, where numerous methodological problems arise, such as missing values in control data source, missing values in the corresponding ESS survey, problems with finding proper population control data (e.g. for non EC countries), specific issues with control data (e.g. regional changes in countries, changes and specifics of education categorization/system etc) and many other similar issues.

The main question here is, whether the weighting will change ESS data, i.e. will the substantive results (i.e. target variables estimates) be different after weighting. Results of extensive earlier studies (Vehovar 2006, Vehovar and Zupanič 2007, Vehovar et al 2010) already showed the following:

- The majority of variables for the majority of countries demonstrated only negligible differences after population weighting for age, gender and education. This is not unusual, because it is a general experience that in surveys with relatively high response rates – which is still true for ESS - the attitude variables rarely demonstrate significant differences after socio-demographic population weighting.
- Sometimes, however, weighting for age-gender-education does result in significant differences (predominantly due to education), particularly with variables related to work, media and politics. The extent of discrepancies (weighted vs. un-weighted data) varies strongly across the countries, with response rate being the strongest correlate.
- Past weighting exercises also detected severe problems with the coding of the education variable and sometimes also with the region variable. This had required

additional clarifications of open issues in past ESS rounds and methodological improvement for 5th round.

We basically expect very similar effects (a, b) also in future rounds of weighting.

Based on the experiences from previous ESS population weighting (Vehovar et al 2010) we first upgraded the inventory of control variables and auxiliary data sources. We proceed with the same variables as in previous attempts: age, gender, education and region (region was added only in some later runs of past weighting procedures).

As regards the education variable, several problems were detected during previous weightings, which were then resolved in collaboration with ESS education experts, who checked and corrected all R1-R4 data and also prepared profound methodological (including coding instructions) improvement in R5. In the most recent ESS data release for R5 thus a much upgraded harmonized education variable is used. Consequently, we further proceed here with preparing a standardized procedure for harmonized weighting, based on new ESS data releases for R1-R4 and fresh R5 data. Similar as in our previous efforts, we limit ourselves only to weighting and we do not involve ourselves with clarifications, which would require checking the primary data, or, contacting specific country official administrations and national coordinators (Appendix C). We assume that the quality of the ESS data was sufficiently resolved before and during the archiving process, so we focus on weighting alone. Nevertheless, we still do comment certain conceptual solution related to data quality, predominantly of ESS coding of the education variable. We should also report that majority of time was still devoted to clarifications of the certain country and round data.

We thus proceed here with preparing a standardized procedure for harmonized weighting. First, the finalized standardized weighting procedure will be applied to Round 5 (2010) of the ESS (October 2012), and then to all previous ESS rounds. The corresponding weights will be then published along with explanatory methodological notes for public users of ESS data (April 2013), as foreseen in DoW (Description of Work) of DACE project, where this is specified as Task 3 (Calculation of improved post-stratification weights to correct for unequal representation for all rounds for all countries) and was further elaborated at the 2011 meeting of the task group (Appendix D).

Our weighting procedure assumes continuous and extensive evaluations of methodology and auxiliary population control data tables. First, the procedure and the data will be evaluated internally, within our institution (UL partner). Next, the procedure will be reviewed within the WP12 and, finally, the quality will be verified and approves at the CST and Methods group level. The process of creating weights for ESS data and the corresponding quality control assurance are presented on Figure 1 (next page). We should say that there is no formal check and loop related to explicit dissemination, contacting, clarification and approval from national coordinators. The existing (revised WP12 Task 3) in DoW (Appendix C) does not foresee systematic comparisons of ESS weighting with national weighting approaches, as this was one of activities eliminated from DoW in the process of negotiations related to the costs reduction of DACE proposal.

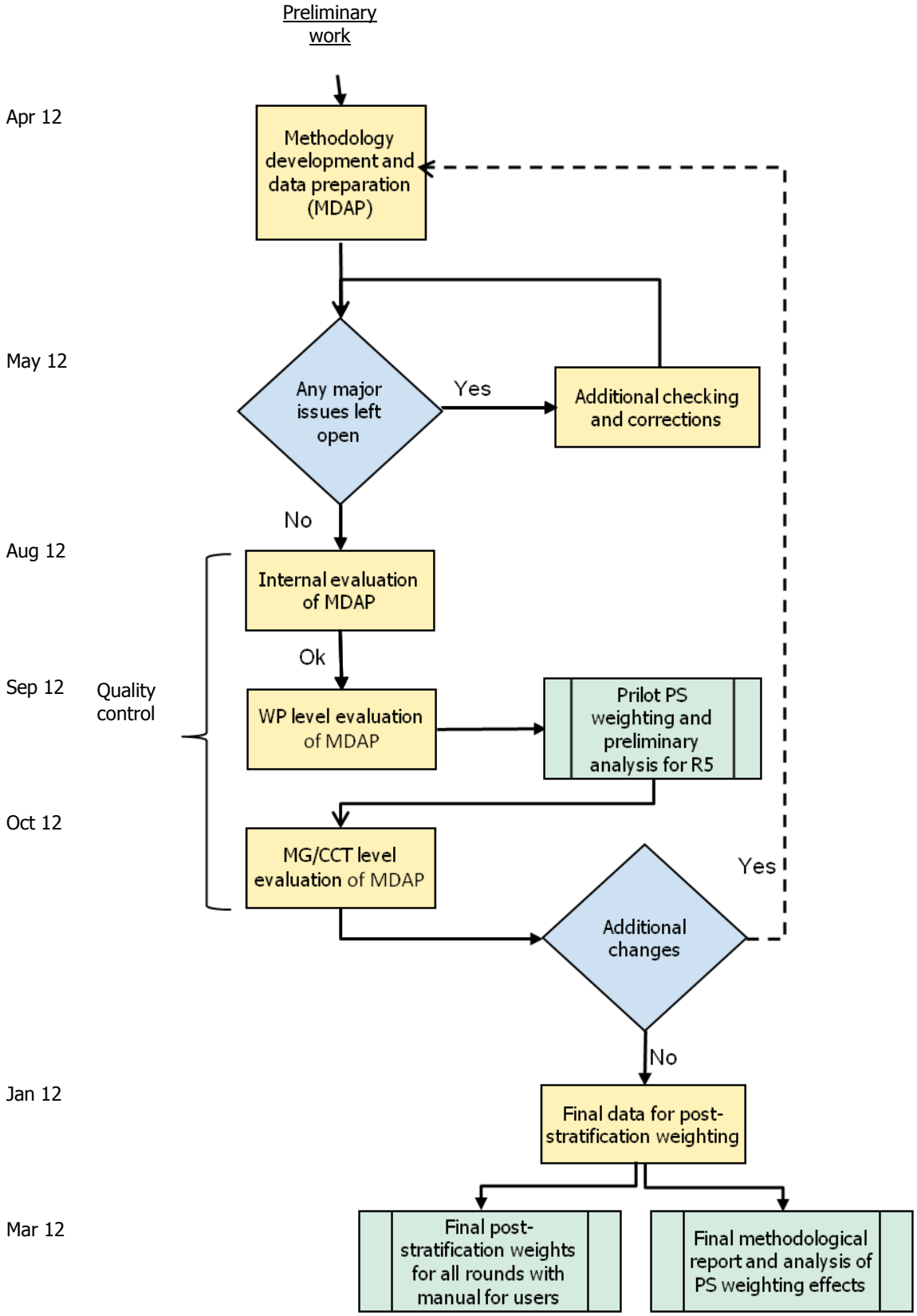
Work on WP13 task 3 is proceeding according to objectives set in the Description of work (Appendix C). Due to additional database cleanings needed there were some delays with corresponding R1-R5 final data releases and additional validations and corrections, which were agreed at the meeting in July 2011 (Appendix D). However, despite this, as mentioned, the deliverables are foreseen to be done on time.

At the very beginning, we should also add a small terminological note: Although this weighting task is formally described in DoW of the DACE project as "post-stratification weighting", we rather use here the word "population weighting", because not only post-stratification but sometimes also raking procedures are used to weight the data.

We also expose that we do not enter here into discussions, whether this is predominantly weighting for nonresponse (which is perhaps still the main reason for discrepancies), as there are also many other sources which lead to the deviations from population controls, from frame deficiencies, sample design simplifications, fieldwork deviations etc.

In certain part we in fact deal here also with so-called "cosmetic weighting" as we correct the data, particularly the margins, to some know external figures, with no or relatively little effect on substantive results, i.e. just to provide users with the feeling of having "true" population margins. However, on the other side, we may in certain cases still expect also more substantive differences, so the weighting is done as a precaution for the deviations, which can otherwise occur if no population weights are implemented.

Figure 1: Process



2 Standardized procedure

2.1 Control data sources

The first challenge when doing post-stratification weighting is the choice of population control data sources. In some cases the basic demographic variables such as gender, age, education and region may not be easily available for every year.

In our case we re-considered potential population statistics provided by national coordinators (ESS Appendix), OECD data (Organization for Economic Cooperation and Development) and LFS data (Labour Force Survey), where both OECD and LFS data use the UNESCO approved ISCED 1997 standard[1], but we decided to keep using LFS data as the basic source (Table 1).

The decision is predominantly due to good and continuous coverage of the data, but also because national LFS teams are typically relatively large and they have the expertise with which they clarify the methodological issues around population controls with the Eurostat.

We should also add that ESS education experts also confirmed that LFS is not ideal and it has certain deficiencies, however, altogether it is still the far best source for our needs.

Table 1: Comparison of population sources

Source	Advantages	Disadvantages
ESS Appendix (Population statistics provided by national coordinators)	Available for all countries	No standardization, not comparable across countries, variation across years, lots of missing values
Organization for Economic Cooperation and Development (OECD)	ISCED97 standard	Does not cover eight ESS countries, less suitable for region (EU NUTS levels)
Labour Force Survey (LFS)	ISCED97 standard, continuous verification of control variables at national statistical offices.	Does not cover three ESS countries

[1] http://www.uis.unesco.org/education/ISCEDmappings/Pages/default.aspx

We should also expose that although we use Eurostat LFS, the LFS data are based on national survey samples, where the control variables are often already nationally weighted (corrected) by proper population values, provided by national statistical offices. Thus, LFS data for control variables provided by Eurostat already include nationally verified population values for control variables and not only the sample estimates. So when sampling variance is considered, the sample variability of the control variables is not an issue, as it is often not sample estimate, so we can still talk about population weighting. However, even in examples, when the national LFS survey data were not weighted for all socio-demographic control data (i.e. by Statistical Offices, before given to Eurostat), the LFS samples are large enough (we use here annual aggregated data), so that the corresponding sampling variability would be negligible.

We should also add that we received the control data from Eurostat very easily, almost instantly, in a form of frequencies in four dimensional (gender, age, education, region) count table.

There are only three countries for which corresponding Eurostat controls form LFS are not available: Israel, Russian Federation and Ukraine. In these cases we use data from the ESS Appendix instead. Moreover, the Appendix data has to be in part used also in Norway, Iceland and Sweden (R5) where we have incomplete data about age (missing category 75+). In addition, we used Appendix data also to improve the LFS population estimates in Croatia (within task 3A). In cases when data are missing only for one round, the neighboring round is used (e.g. Turkey). Exact sources of the data are listed in Table 2.

Table 2: Source of control data and eventual corrections done in ESS sample data

	R1	R2	R3	R4	R5
AT	LFS	LFS	LFS	-	-
BE	LFS	LFS	LFS	LFS	LFS
BG	-	-	LFS	LFS	LFS
CH	LFS	LFS	LFS	LFS	LFS (2013)
CY	-	-	LFS	LFS	LFS
CZ	LFS	LFS	-	LFS	LFS
DE	LFS	LFS	LFS	LFS	LFS
DK	LFS	LFS	LFS	LFS	LFS
EE	-	LFS	LFS	LFS	LFS
ES	LFS	LFS	LFS	LFS	LFS
FI	**LFS for GAE, APP for region**	**LFS for GAE, APP for region**	**LFS for GAE, APP for region**	**LFS for GAE, APP for region**	**LFS for GAE, APP for region**
FR	LFS	LFS	LFS	LFS	LFS
GR	-	LFS	-	LFS	LFS
HR	-	-	-	**LFS adjusted with APP**	**LFS adjusted with APP**
HU	LFS	LFS	LFS	LFS	LFS
IL	**APP**	-	-	**APP**	**APP**
IS	-	LFS	-	**LFS APP (75+)**	-
IE	LFS	LFS	LFS	LFS	LFS
IT	LFS	LFS	-	-	-
LU	LFS	LFS	-	-	-
LT	-	-	-	**LFS (EDU ADJ)**	-
LV	-	-	LFS	LFS	-
NL	LFS	LFS	LFS	LFS	LFS
NO	**LFS APP (75+)**	**LFS APP (75+)**	**LFS APP (75+)**	**LFS APP (75+)**	**LFS APP (75+)**
PL	LFS	LFS	LFS	LFS	**LFS**
PT	LFS	LFS	LFS	LFS	LFS
RO	-	-	LFS	LFS	-
RF	-	-	**APP (EDU ADJ)**	**APP (EDU ADJ)**	**APP (EDU ADJ)**
SE	LFS	LFS	LFS	LFS	**LFS APP (75+)**
SI	LFS	LFS	LFS	LFS	LFS
SK	-	LFS	LFS	LFS	LFS
TR	-	LFS (R4)	-	LFS	-
UA	-	**APP**	**APP**	**APP**	**APP**
UK	**LFS**	**LFS**	**LFS**	**LFS**	**LFS**

- = Country did not participate in that round

LFS/APP = Source used for control data (Census Bureau of Statistics 2002 in Israel, Unknown source in Norway R1 Appendix, State Census 2002 in Russian federation and Census 2001 in Ukraine) [in all rounds]

LFS (2013) = LFS data received in June 2013

(75+) = LFS used in general, except for data misses age category (75+) (see section 2.2.1)

(**EDU ADJ)** = Sample data needed adjustment to match control data (see section 2.2.2)

(R4) = No data for certain for particular round, neighboring round used instead

2.2 Control variables

2.2.1 Gender and age

Gender ("gndr" in ESS datasets) is recoded to "gndrR" as: 0 Missing, 1 Male, and 2 Female. This variable is the one with least complications. There are always two categories, male and female, although sometimes we could find missing values when using two two-dimensional tables with gender as one variable and age or education as the second one.

Age ("agea") is recoded to "ageR" as: 0 missing, 1 15-34 years old, 2 35-54 years old, and 3 55+ years old. There is a small issue with age in LFS where data for age groups above 75 years is not provided, i.e. Iceland and Norway. We addressed the problem of missing population data for older population by incorporating control data (i.e. margin for age for 75+) from the ESS appendix, which has complete data for age. However, with that we then lack full interaction with education. Thus, a modified weighting approach will be used for these two countries (see section 2.4).

2.2.2 Education

The **education** variable ("edulvla") in ESS data is recoded to "edulvlvR":

0. Missing;
1. Lower education (lower secondary or less) includes ISCED "level 0 Not completed primary education", "1 Primary or first stage of basic", and "2 Lower secondary or Second stage of basic education". We thus talk here about education, which usually ends between the age of 14 and 16. In principle, according to recent ESS educational instructions, this should also include short vocational programs (less than 3 years) taken after primary school (shorter 3C programs), labeled in LFS with "22".
2. Medium education (higher secondary and post-secondary, non-tertiary) includes ISCED level "3 Upper secondary (A, B, C)" and "4 Post-secondary, non-tertiary". It usually ends between the age of 18 and 20, depending on the length of the programs. Note that vocational programs shorter than three years ("22", which are still in ISCED 3C) are excluded here, as we elaborated above. However, in practice this is not always the case nor it is easy to achieve, as ISCED 3C may stand as one segment in LFS or ESS.

3. Higher education (post-secondary) includes ISCED level 5 and higher levels, i.e. any stage of tertiary education (BA, BSc, MA, PhD …), including vocational ISCED 5B programs which have different names in different countries.

The recoding in ESS is basically analogue to the recoding of LFS data[2] based on the UNESCO ISCED 1997 classification.[3] However there are, as mentioned certain ambiguities around ISCED level 3C, which is difficult to capture and even more difficult to split among the Low (where "22" part of 3C need to be) and Medium level. Thus, in practice, the recoding is often different from the desired classification (Table 3). For instance, by putting the entire ISCED 3C into Medium level, the persons with lower educational education (i.e. 1-2 years of vocational training, category "22") are merged with those with upper secondary education (e.g. finishing gymnasium), because they could not be easily separated from those with full 3 year education of ISCED 3C level. This is the situation for example with Slovenia in ESS and also in LFS[4] (Appendix B), where there is no category "22" in LFS. However, we should add here that in this specific case the Slovenian national education experts also agreed that due to small size of this segment the existing categorization is still acceptable. In any case, this example shows, that the reality, which can be achieved, is often different to formally (officially) declared methodology for ESS and LFS as outlined in Table 3.

Thus, we have a specific technical problem with LFS (and consequently with ESS) – i.e. not being always able to separate category 3C short "22" from remaining ISCED 3C long (code "31") and put it into Lower level. The ESS and LFS thus often unintentionally (not being able to prevent such deviations) allow the shorter ISCED 3C (LFS code "22") into the Medium level. In addition, we should also point to another conceptual issue with education, which is however worth some discussion, at least for social scientists. Namely, the 3 (and more) year upper secondary vocational education - which does not lead to tertiary education (i.e. ISCED

[2] Levels of education and training in the EU Labour Force Survey:
http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/ISCED_EN.htm

[3] ISCED 1997: http://www.uis.unesco.org/education/ISCEDmappings/Pages/default.aspx

[4] The category of less than three years of 3C (labeled in LFS as »22« does not exsist in LFS data for Slovenia (see Appendix). Obviously, they were merged with LFS code »31« (3C), so it is not possible to separate this group (1-2 year vocational) from the Medium and shift them to Low education. To do this, we would need to contact national statistical office.. Similarly, in ESS they are not separated (they ara all in one joined ISCED 3C category) and further investigation would be needed to separate them. We should add that category »22« at LFS can be found only in AT, CH, CY, DK, ES, GR, IE, IS, IT, LU, MT, SE in UK. More>>.

3C, LFS code "31") - is merged in LFS and ESS with other upper secondary categories (ISCED 3A and ISCED 3B) into Medium level. (As discussed above also the code "22" of ISCED 3C is sometimes joined here in many countries, what is of course even more problematic). This may be somehow questionable from the content points of view. Very often, even according to marketing research categorization, persons with completed 4-5 year secondary education (3A or 3B) , which may lead to tertiary education (ISCED 5) differ a lot (due to different attitudes, lifestyles, spending etc.) from those with finished (vocational) upper secondary education which does not lead to tertiary levels (3C). Thus, it would be perhaps better to (a) merge the entire 3C (not only 1-2 year program of 3C) with the Lower educated or (b) have them – either entire 3C, or, only 3C code "31" (without "22" which may remain in "Lower level") as a separate (fourth) educational category, because these are usually very substantial segments of the population. All 1-3 year vocational educations/training (3C) is very often amounting around a quarter of the target population. However, given the available LFS data, this would entirely change the data preparation (at ESS and LFS side) and also the corresponding weighting strategy, so this would require some extra efforts and resources. It is also true, that in some countries (e.g. Germany), the 3C and 3B are in fact very much similar, because 3B rarely go to ISCED 5 level, so 3B in much closer to 3C than to 3A, so at least option (a) would not be best solution, as it would split relatively homogenous segment among Low (3C) and Medium (3B).

In Table 3 both classifications are structured and compared. The LFS simply follows the official ISCED mappings, where the three-level structure is then imposed by official LFS structure. As mentioned, there is already certain controversy in LFS about shorter (1-2 year) 3C programs (LFS code "22"), as they are actually not always successfully separated and included into Low level. In most countries, these differences are small and negligible, while in the case of serious discrepancies (i.e. England, Poland) the ESS education experts intervene and modify this classification, so there the ESS now departs from LFS (we elaborate this further below).

With respect to ESS, there is no initial and official ESS separation into Low, Medium and High level of education as it is the case with ESS. We introduce this structure here only for the weighting task, where we try to follow the LFS as control data source.

As there were for past rounds only six aggregated ESS categories of the education (i.e. variable "edulvla"), which can be harmonized for all ESS surveys, we proposed the merging as in Table 3 below:

- 0, 1, 2 to form Low level;
- 3, 4 to form Medium level,
- 5, 6 to form High level.

The lightly shaded areas in Table 3 denote the problematic ISCED 3C (code "22") segment, which is in practice often allocated into wrong level.

Table 3: Target education categories and education recoding in ESS and LFS data

	ESS (EDULVLA variable)	Low		Medium		High
	Official LFS (HATLEVEL variable)	0-1 Not completed primary; Primary or first stage of basic education	2 Lower secondary or second stage of basic education	3 Upper secondary education	4 Post secondary, non-tertiary education	5 -6 First stage of tertiary; Second stage of tertiary education
Low	ISCED 0 — Pre-Primary Education					
	ISCED 1 — Primary Education Or First Stage Of Basic Education					
	ISCED 2 — Lower Secondary Education Or Second Stage Of Basic Education					
	ISCED 3C: Programs not designed to lead to ISCED 5A or 5B (shorter than 3 years).					
Medium	ISCED 3C: Programs not designed to lead to ISCED 5A or 5B (3 years or more).					
	ISCED 3 — (Upper) Secondary Education					
	ISCED 3A: Programs designed to provide direct access to ISCED 5A;					
	ISCED 3B: Programs designed to provide direct access to ISCED 5B;					
	ISCED 4 Post-Secondary Non Tertiary Education					
	ISCED 4A, 4B, 4C: See text for ISCED 3					
High	ISCED 5 — First Stage Of Tertiary Education (Not Leading Directly To An Advanced Research Qualification)					
	ISCED 5A, ISCED 5B					
	ISCED 6 — SECOND Stage Of Tertiary Education (Leading To An Advanced Research Qualification)					

Both of the above problems are in large part beyond decisions we can make here in this narrow weighting exercise, because for both issues - i.e. putting the 1-2 year programs from 3C (code "22") into Lower level and separating the entire 3C from 3A and 3B – may require extensive contacting with country administration and further research with primary control data. As mentioned above, here our work (Appendix C) is strictly limited to the weighting issues. At most we perform desk research, consistency checks and consultations to ESS education experts, but we do not enter into evaluation of data quality (e.g. LFS or ESS appendix data) nor do we go against the decisions of official ESS education coding or against the decisions of ESS education experts.

Besides above two major issues with education we also have some more specific problems. First, there are three countries for which there are no LFS data (Israel, Russia and Ukraine) and, second, there are two countries without LFS data in one of the rounds (Switzerland and Turkey). For Switzerland we used Round 4 data as a control for weighting Round 5 and for Turkey data from Round 3 was used to weight Round 2, while for Israel, Russia and Ukraine we used population data provided by national ESS coordinators (ESS Appendices). We should remind here again that we accepted these data, as our role is to perform weightings and not to extensively validate the data sources (i.e. ESS Appendix reports).

It turns out that when applying the above-described grouping into three categories (Low, Medium, High), ESS and LFS frequencies are comparable in most of the countries. However, there were three countries with major discrepancies: Lithuania (R4), Poland (R5) and United Kingdom (R1-R5). The mismatch originates from the decision, done by ESS educational experts, to change the recoding of the variable to make educational distributions more comparable across ESS countries. According to ESS educational experts the official data in certain countries is distorted. For instance in UK, it shows less low educated and more medium educated than there actually are. Thus, the ESS changed certain UK recoding of ESS which resulted in deviations from official ISCED mappings used in LFS.

On the other hand, from our weighting strategy perspective, as elaborated and justified in previous reports (Vehovar et al 2008), a closer match with official data (LFS) should have precedence over inter-country comparability. In order to resolve the mismatch between ESS and LFS data, and on explicit advice and accordance with ESS educational experts, for the purposes of developing PS weights, the recoding in the ESS datasets in Lithuania (R4),

Poland (R5) and United Kingdom (R1-R5) was changed to official ISCED mappings to regain comparability with control data. The effects of this intervention are described in proceeding text and tables. We should expose that this was done only for the purpose of weighting (so to match LFS) and such categorization will not be used for ESS users.

Let us closely look at these four exceptions, starting with the United Kingdom where the lower educated make up 28% of the ESS sample in 2002 according to official mappings, because, originally the GSCSE/O-lovel/NVQ1/NVQ2 was coded as upper secondary (ISCED 3C) by the ESS, and put into Medium level. After that ESS educational experts intentionally recoded it to lower secondary (ISCED2), in order to improve inter-country comparisons, as these segment include 1-2 year programs (but also other types of programs). This change increased the lower educated segment to 55% which visibly deviates from the 24% of lower educated in the LFS. To match the ESS with the LFS educational variable, following the suggestion of ESS educational experts, we recoded it back again to upper secondary (ISCED3). The re-adjustment is done for UK in all rounds. In R5 an additional correction was made for the ISCED 4a and 4b which were recoded to ISCED 5 to make data more comparable to previous years and other European countries.

Similarly, in Poland (R5) the short ISCED 3C programs (Vocational, less than two years, no access to ISCED 5 level) were initially coded into Medium, but were later recoded to Lower education (together with ISCED 2 and lower) by ESS educational experts because of recent educational reforms not taken into account in LFS data where all respondents are treated as if they finished their education after the reforms although the opposite is closer to reality. To match LFS data we recoded it to lower tier upper secondary (ISCED 3), although we believe leaving it in the lower category would be more appropriate and that there is in fact a coding mistake (or insufficient elaboration) in LFS data. In fact, the situation in Poland is very similar to Slovenia (where this intervention was not done): the short 3C are merged with longer 3C into Medium level, so it is now not possible to separate them and allocate to Low and Medium correspondingly. Thus, we proceed in Poland with the adjustment which decreased the lower education category from 44 back to 22% so that it matches the 23% rate in LFS data. In previous rounds these adjustments of the recoding were not needed.

Interventions from ESS educational experts in UK and Poland are very much in line with our comments regarding the general approach in LFS and ESS coding (at least in R1-R4) scheme,

which rather unreasonably coded secondary programs that are not designed to lead to tertiary education into the same level (Medium) as programs which lead to tertiary programs (3A, 3B). Unfortunately, we could not accept this approach – despite the fact it is right - in some countries, because this would destroy the comparability and consistency of using LFS as the control.

The ESS educational expert interventions are to be discussed seriously also in the future, as new ESS coding instructions for R5 establish a completely new platform for educational weighting. If the separation of 3C would be now feasible for all ESS countries, then we should consider further elaboration also for LFS data, which may in principle provide sufficient desegregation, but this would require extra contacts with countries, because LFS receives only aggregation of national education variable.

However, at this point we can establish that ESS education interventions are not fully consistent for the ESS, not even in R5, as interventions done in Poland and UK[5] were not done in all countries, but it is estimated that in other countries this is of relatively limited impact. As R5 education coding instruction in principle enables required desegregation, we may even think about discontinuity in weighting procedure, because for past rounds it is very difficult to get more details.

On the other hand, there are also issues in Lithuania, about the boundary between medium and high education. In Lithuania (R4) the ESS educational experts recommended coding Special secondary education as Tertiary education so that it would be better comparable with other Baltic and Eastern European countries. To match control data the recoding was changed to post-secondary non-tertiary education. The higher educated that made up 42% of the sample before adjusting decreased to 25% to match the 23% higher educated in LFS data. In other rounds the recoding was not needed.

[5] Finally, after consulting national coordinators (task 3A) the education variable was not used to weight UK and PL were not used in the weighting procedure.

Table 4: Adjusted recoding for United Kingdom, Poland, and Lithuania

Data			isced 0-1	isced 2	isced 3	isced 4	Isced 5-6	othr	miss	total	L	M	H
UK	2002	ESS	28.29	26.65	13.28	0.00	31.51	0.00	0.27	100	**55.09**	**13.32**	**31.59**
		ESS adj	28.29	0.00	39.93	0.00	31.51	0.00	0.27	100	**28.37**	**40.04**	**31.59**
		LFS	0.20	24.12	28.38	0.00	18.64	0.00	28.66	100	**34.09**	**39.78**	**26.12**
	2004	ESS	31.53	9.42	18.05	0.00	34.92	5.78	0.30	100	**43.60**	**19.21**	**37.18**
		ESS adj	31.53	0.00	27.47	0.00	34.92	5.78	0.30	100	**33.57**	**29.24**	**37.18**
		LFS	0.15	23.43	33.91	0.04	20.21	0.00	22.26	100	**30.33**	**43.67**	**25.99**
	2006	ESS	23.70	23.22	13.08	0.00	39.59	0.00	0.40	100	**47.11**	**13.13**	**39.75**
		ESS adj	23.70	0.00	36.30	0.00	39.59	0.00	0.40	100	**23.79**	**36.45**	**39.75**
		LFS	0.22	21.56	34.85	0.09	21.13	0.00	22.16	100	**27.98**	**44.88**	**27.15**
	2008	ESS	23.40	19.10	12.96	0.00	44.02	0.00	0.52	100	**42.73**	**13.03**	**44.25**
		ESS adj	23.40	0.00	32.06	0.00	44.02	0.00	0.52	100	**23.53**	**32.22**	**44.25**
		LFS	0.32	26.15	34.84	0.05	23.08	0.00	15.56	100	**31.36**	**41.31**	**27.33**
	2010	ESS	22.48	16.87	25.69	2.77	28.91	2.58	0.70	100	**40.68**	**29.43**	**29.88**
		ESS adj	22.48	0.00	42.56	0.00	31.68	2.58	0.70	100	**23.24**	**44.00**	**32.75**
		LFS	0.30	23.46	34.99	0.06	25.23	0.00	15.96	100	**28.28**	**41.70**	**30.02**
PL	2010	ESS	3.61	40.29	31.16	3.76	20.90	0.00	0.28	100	**44.03**	**35.02**	**20.96**
		ESS adj	3.61	18.76	52.69	3.76	20.90	0.00	0.28	100	**22.43**	**56.61**	**20.96**
		LFS	4.05	19.31	55.53	2.97	18.14	0.00	0.00	100	**23.36**	**58.51**	**18.14**
LT	2008	ESS	11.44	16.38	24.88	5.59	41.71	0.00	0.00	100	**27.82**	**30.47**	**41.71**
		ESS adj	11.44	16.38	24.88	22.43	24.88	0.00	0.00	100	**27.82**	**47.30**	**24.88**
		LFS	12.18	14.49	31.95	18.37	23.01	0.00	0.00	100	**26.67**	**50.32**	**23.01**

Certain pre-weighting corrections in educational recoding were needed also for Israel (R1, R4), Russia (R3, R4 and R5) and Ukraine (R2-R5), where ESS Appendix was used as a source of population controls (see Table 5 on next page).

In Israel (R1 and R4) the adjustment was performed by recoding the category "Post-secondary, non-tertiary" to ISCED 4 and grouping it as medium, not higher. In R1 it included also correcting the recoding of full vocational schools to ISCED 3 and grouping them as medium education, not lower. For instance, there were originally 35% lower, 28% medium and 37% higher educated in the ESS sample in 2002. After the adjustment the structure changed to 22% lower, 53% medium and 25% higher educated which better corresponded to the structure in LFS data: 19% lower, 57% medium and 24% higher.

In Ukraine (R2-R5) the adjustment consisted of recoding the category "Secondary technical education" (ISCED 5B short, advanced vocational qualifications) to ISCED 4 and groping it

as medium, not higher education. For instance, there were 51% highly educated in the ESS sample in 2004. The adjustment decreased them to 15% to be closer to the 14% in LFS data[6].

In Russia (R3-R5) the adjustment was recoding the advanced vocational education (ISCED 5B short, advanced vocational qualifications) to ISCED 4 and then grouping it as medium, not higher education. For instance, in 2006 the 48% highly educated according to ESS data were adjusted to 20% to match the 19% in LFS data.

Table 5: Adjusted recoding for Ukraine, Russia and Israel

Data			isced 0-1	isced 2	isced 3	isced 4	Isced 5-6	othr	miss	total	**L**	**M**	**H**
UA	2004	ESS	12.12	8.70	28.40	0.00	50.41	0.00	0.36	100	**20.90**	**28.51**	**50.59**
		ESS adj	12.12	8.70	28.40	35.27	15.14	0.00	0.36	100	**20.90**	**63.90**	**15.20**
		ESS app	-	-	-	-	-	-	-	-	**27.87**	**58.02**	**14.11**
	2006	ESS	11.64	8.91	26.89	0.00	52.52	0.00	0.03	100	**20.56**	**26.90**	**52.54**
		ESS adj	11.64	8.91	26.89	34.73	17.79	0.00	0.03	100	**20.56**	**61.65**	**17.79**
		ESS app	-	-	-	-	-	-	-	-	**27.87**	**58.02**	**14.11**
	2008	ESS	4.39	9.81	17.58	12.86	55.03	0.00	0.33	100	**14.24**	**30.55**	**55.21**
		ESS adj	4.39	9.81	27.31	29.12	29.04	0.00	0.33	100	**14.24**	**56.62**	**29.13**
		ESS app	-	-	-	-	-	-	-	-	**27.87**	**58.02**	**14.11**
	2010	ESS	3.31	7.00	22.04	8.88	58.64	0.06	0.07	100	**10.32**	**30.96**	**58.72**
		ESS adj	3.31	7.00	22.04	33.70	33.82	0.06	0.07	100	**10.32**	**55.82**	**33.86**
		ESS app	-	-	-	-	-	-	-	-	**27.87**	**58.02**	**14.11**
RF	2006	ESS	6.72	12.50	32.90	0.00	47.88	0.00	0.00	100	**19.22**	**32.90**	**47.88**
		ESS adj	6.72	12.50	32.90	28.23	19.65	0.00	0.00	100	**19.22**	**61.12**	**19.65**
		ESS app	-	-	-	-	-	-	-	-	**22.72**	**58.01**	**19.28**
	2008	ESS	5.66	10.04	31.71	0.00	52.60	0.00	0.00	100	**15.69**	**31.71**	**52.60**
		ESS adj	5.66	10.04	31.71	27.72	24.87	0.00	0.00	100	**15.69**	**59.43**	**24.87**
		ESS app	-	-	-	-	-	-	-	-	**22.72**	**58.01**	**19.28**
	2010	ESS	3.25	9.65	34.83	0.00	52.28	0.00	0.00	100	**12.90**	**34.83**	**52.28**
		ESS adj	3.25	9.65	34.83	25.62	26.65	0.00	0.00	100	**12.90**	**60.45**	**26.65**
		ESS app	-	-	-	-	-	-	-	-	**22.72**	**58.01**	**19.28**
IL	2002	ESS	1.23	33.37	27.46	0.00	36.50	0.00	1.45	100	**35.10**	**27.86**	**37.04**
		ESS adj	1.23	20.02	32.56	20.01	24.74	0.00	1.45	100	**21.56**	**53.34**	**25.10**
		ESS app	-	-	-	-	-	-	-	-	**19.43**	**56.65**	**23.91**
	2008	ESS	2.60	10.36	45.04	0.00	41.49	0.00	0.51	100	**13.03**	**45.27**	**41.70**
		ESS adj	2.60	10.36	45.04	14.82	26.67	0.00	0.51	100	**13.03**	**60.16**	**26.81**
		ESS app	-	-	-	-	-	-	-	-	**19.43**	**56.65**	**23.91**

* ESS = data in ESS database, ESS adj = ESS data with adjusted education recoding, LFS = control data from labour force survey, ESS app = control data from ESS appendix

[6] Finally, after consulting national coordinators (task 3A) the education variable was not used to weight UA.

2.2.3 **Region**

Region data in LFS files is available for all countries, except for Israel, Russia and Ukraine. For Israel and Russia we used data in the ESS appendices that was provided by the national coordinators, while for Ukraine we used census data from 2011 (State Statistics Committee of Ukraine). In LFS we have missing data also for Denmark 2002. In this case data for 2004 will be used as a control.

In contrast to education which is standardized to three levels, each country has a different region variable which varies in the number of categories. All control data (LFS) are given at NUTS2 level (Eurostat NUTS), while some countries in the ESS use different classifications: NUTS1 (less detailed), NUTS2, NUTS3 (more detailed) or sometimes even partially aggregated NUTS2 or NUTS3 classifications (Switzerland, Greece, Portugal and Ukraine). Moreover, some countries use a different classification in different rounds, in particular Round 5 differs from other rounds since an improved methodology was applied that year where data are often more detailed and complete (note that also variable names were changed). For weighting purposes the region data is recoded to common denominator so that ESS and LFS categories match as presented by Figure 2 and Table 3.

When control data (NUTS2 level) has more categories than ESS data, the former is usually recoded to NUTS1 level (Belgium, Germany, France, Turkey and UK), but in some instances to another common denominator, which is actually a partial aggregation of NUTS2 into a lower number of categories. This is needed when one of the rounds has a different number of categories than others (Switzerland, Greece, Finland), or, when some regions are excluded (Portugal, France), or a non-NUTS coding used in population data (Ukraine). In some of the countries the recoding is not needed in all rounds, namely Greece (R1, R2, R3 and R5), Switzerland (R2, R3, R4 and R5), Belgium (R5) and France (R5) where the sample and control data already match in the number of categories for region.

When ESS sample data is more detailed (usually NUTS3) than LFS control data (NUTS2), then the former is recoded to NUTS2 level. This is the case in Bulgaria, Czech Republic, Cyprus, Estonia, Ireland Lithuania, Latvia, Netherland, Slovenia, Slovakia and Denmark. In Denmark (R4) and in Czech Republic (R4 and R5) the recoding is actually not needed as the sample and control data already match. Note also that in Cyprus, Estonia, and Lithuania and

Latvia data the common denominator is one category and so data will not be weighted by region. Of course, with this intervention we lose information, but NUTS3 level inclusion would require substantial changes – which is one of the options to be discussed - in our approach, as LFS includes only NUTS2.

If sample and control data have the same number of categories, usually no recoding is needed (Croatia, Israel, Iceland, Luxembourg, Norway, Romania, Russia), but sometimes the labels or order of precedence have to be changed (Austria, Poland, Hungary, Sweden). Note that Iceland and Luxembourg have only one category and data will not be weighted by region. In Hungary and Sweden there is an exception in Round 5 where NUTS1 is used on ESS data so it has to be recoded to NUTS2 to match the classification used in other rounds and in control data. Another exception is Spain and Italy, where sample data have some empty cells. In these two cases data is recoded to NUTS1 (less detailed) as a common denominator.

To sum up, the common denominator can be:

1. **NUTS2** – ESS values (NUTS 3 level) are recoded to this level (e.g. Bulgaria); or no recoding is needed if ESS is already in NUTS2 (e.g. Austria)
2. **NUTS1** – LFS values (and if needed also ESS values) are recoded to this level (e.g. Germany). In addition, also in some cases when ESS regions are given at NUTS2 level we recode to NUTS1 because of incomplete data (e.g. Spain).
3. **Other denominators** (e.g. partially aggregated NUTS2) – both ESS and LFS values are recoded (e.g. Finland from 5 to 4 categories). Another denominator has to be used in specific cases in one of the rounds of (e.g. Switzerland).

Again, note that in Iceland and Luxembourg (no recoding needed) and Cyprus, Estonia, Lithuania and Latvia (LFS recoded from NUTS2 to NUTS1) the common denominator is only one region. In these countries weighting by region will not be performed at all. Here again we face the dilemma whether we should look for more detailed data at NUTS3 level or maybe even country specific regional classifications. As mentioned, this would require extra investigation into the control variables and substantially more resources.

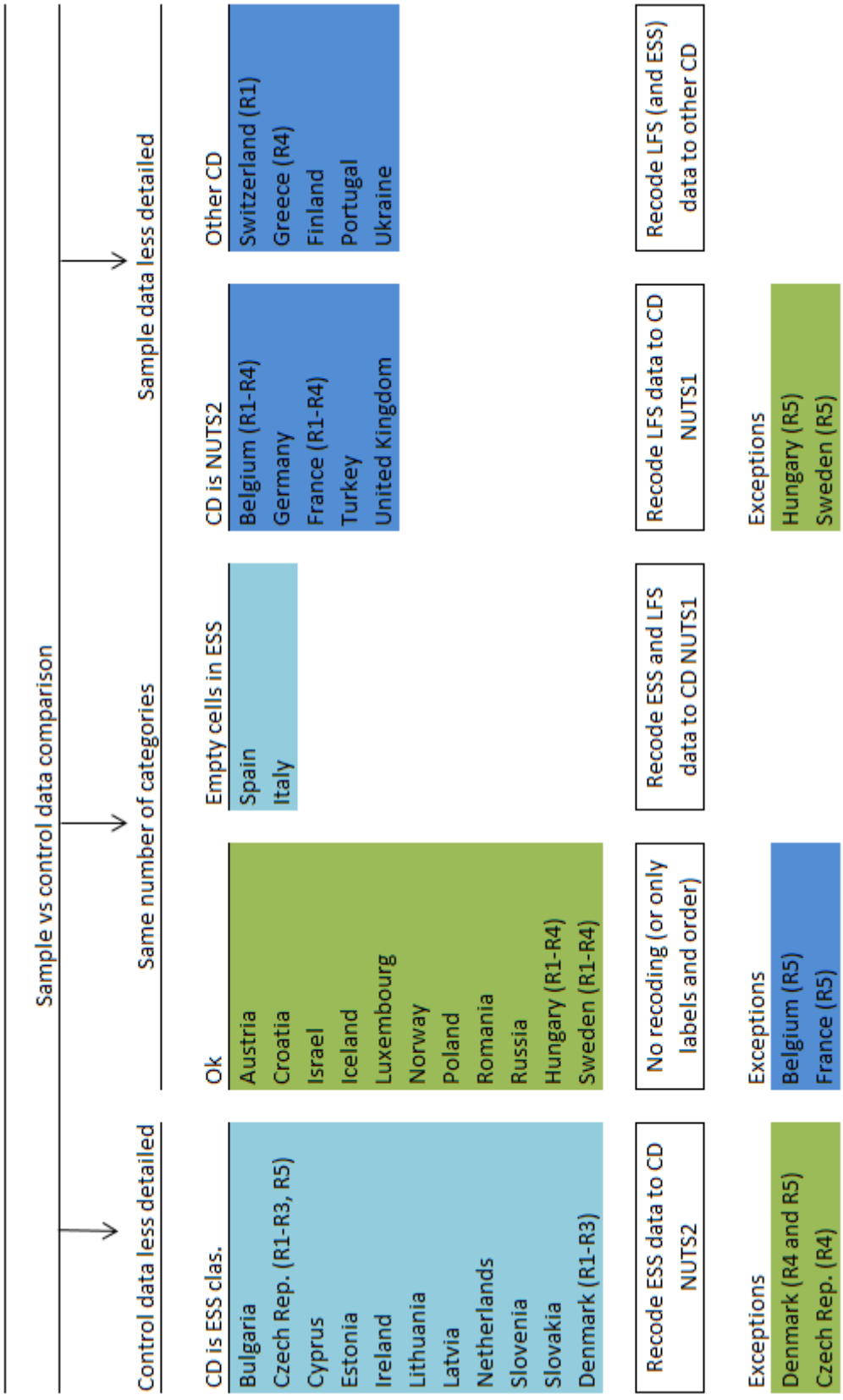
At control data side, there are sometimes missing values for region, as in Czech Republic (0.3%) and Sweden (10%). In the first case they are very small and easily ignored. In Sweden, however, there are a lot of missing values but the distribution of known control values is in

full correspondence with the one on the sample, so we assume there is not much harm in ignoring them (treating them as missing completely at random, MCAR). On the other hand, Slovakia has missing values on the sample (up to 1%) which was resolved by inserting the missing cell in control data and evenly decreasing other values. The changes are fully documented in our appendices.

Table 6: Region review

	ESS	ESS	NUTS2	RECODE	Notes, exceptions, etc.
AT	NUTS2	9	9	/	Only change order of precedence.
BE	NUTS1	3 - 11	11	NUTS2 to NUTS1.	R5 matches NUTS2 but we still recode it to NUTS1.
BG	NUTS3	28	6	NUTS3 to NUTS2.	
CH	NUTS2	7	7	NUTS2 to 6 regions.	Only R1 has 6 regions (later rounds match NUTS2) but the recoding is done for all for better comparison.
CY	NUTS3	9	1	NUTS3 to NUTS2.	
CZ	NUTS3	8 - 14	8	NUTS3 to NUTS2.	Except for R4 when NUTS2 is already used in ESS.
DE	NUTS1	16	33-34	NUTS2 to NUTS1.	Note that in R1 there were 33 regions in NUTS2.
DK	NUTS3	5 - 15	5	NUTS3 to NUTS2.	Except for R4 and R5 when NUTS2 is already used in ESS. There is no LFS data for R1, R2, and R3 so we use R4 data.
EE	NUTS3	5	1	NUTS3 to NUTS2.	
ES	NUTS2*	18	16	NUTS2 to NUTS1.	16 regions because of certain empty cells in ESS region data.
FI	NUTS3*	4 - 19	5	NUTS2 to 4 regions.	R1 (5 regions but not NUTS2) and R5 (19 regions, NUTS3) are also recoded to 4 regions (Census data used instead of LFS).
FR	NUTS1	9 - 25	25	NUTS2 to NUTS1.	R5 matches NUTS2 but we still recode it to NUTS1.
GR	NUTS2*	10 - 13	13	NUTS2 to 10 regions.	Because in R4 we have only 10 regions we recode also all other rounds to 10 regions.
HR	NUTS2	3	3	/	Weighting possible without recoding.
HU	NUTS2	7 - 20	7	NUTS3 to NUTS2.	Recoding needed only for R5 (NUTS3, 20 regions). For others only change order of precedence and labels.
IE	NUTS3	3 - 8	2	NUTS3 to NUTS2.	Different number of regions in each round (3 in R3, 4 in R4, 8 in others). Recode all to NUTS2.
IL	non-NUTS	7	7	/	Weighting possible without recoding.
IS	NUTS2	1	1	/	No recoding, no weighting with region.
IT	NUTS2*	20	21	NUTS2 to NUTS1.	We recode to NUTS1 (5 regions) because of certain empty cells in ESS region data.
LT	NUTS3	10	1	NUTS3 to NUTS2.	
LU	NUTS2	1	1	/	No recoding, no weighting with region.
LV	NUTS3	6	1	NUTS3 to NUTS2.	
NL	NUTS3	40	12	NUTS3 to NUTS2.	
NO	NUTS2	7	7	/	Weighting possible without recoding.
PL	NUTS2	16	16	/	Only change order of precedence.
PT	NUTS2*	5	7	NUTS2 to 5 regions.	In ESS two island regions (Azores & Madeira) are excluded. Change also order of precedence.
RO	NUTS2	8	8	/	Weighting possible without recoding.
RU	non-NUTS	10	10	/	Weighting possible without recoding.
SE	NUTS2	8 - 21	8	NUTS3 to NUTS2.	Recoding needed only for R5 (NUTS3, 21 regions). For others only change order of precedence.
SI	NUTS3	12 - 16	2	NUTS3 to NUTS2.	Note that in R5 a new classification is used with 16 regions.
SK	NUTS3	8	4	NUTS3 to NUTS2.	
TR	NUTS1	12	26	NUTS2 to NUTS1.	
UA	non-NUTS	22 - 24	27	non-NUTS to 11 regions	We recode to NUTS1 (11 regions) because of certain empty cells in ESS region data.
UK	NUTS1	12	37	NUTS2 to NUTS1.	

Figure 2: Region review



2.3 Strategy for dealing with missing data

Control variables, especially education, can have a lot of missing values on sample and on control data. This is an issue, particularly when preparing tri-dimensional post-stratification tables of gender, age and education (GAE).

For each cell in the post-stratification table, there are three situations how missing data can appear:

1. **Missing value only in sample cell**. Values in missing cells are copied to corresponding cells in control table, taking the missing at random assumption (MAR). Next, other cells in the control table are proportionally adjusted so that the total sum and the ratio between existing cells are preserved.

2. **Missing value only in control data**. Usually, we ignore them, using the so called missing completely at random assumption (MCAR). However, if unknown values present more than 1 % of the population (i.e. *DK in R1, EE in R3 and R5, ES in R5, HU in R2, R4 and R5, LV in R4*), assuming MCAR is risky as population structure could be affected. In these cases there is another alternative, assuming an equal distribution of unknown values, i.e. missing at random (MAR), and equally re-allocating them between known values.

3. **Missing values both in sample and in control cell**. If the missing value in sample is lower than in the control, then the cell is normally used in post-stratification. On the other hand, cells where the missing value on control data is substantially higher than on sample data (i.e. *DE R1 and R2, DK in R4 and R5, EE in R2 and R4, HU in R1 and R3, LV in R3, No in R1, R5 and R5, SE in R1-R4, SI in R1*) are treated in a similar way as in cases with missing value only in control data (item 2 above). The control value was decreased to the sample value and the equally re-allocated among other values assuming missing at random (MAR).

A review of situations in all countries and rounds is presented in Table 7.

Table 7: Review of missing value cells in ESS and control data when GAE table is used

	R1	R2	R3	R4	R5
AT	S	S+(P)	S	-	-
BE	NoE	NoE	NoE	NoE	NoE
BG	-	-	S	N	S
CH	NoE	NoE	NoE	NoE	NoE
CY	-	-	S	S	S
CZ	S+(P)	S	-	(P)	(P)
DE	S+P *	S+P *	S	S+(P)	S+(P)
DK	S+P	S+(P)	N	(P) *	P *
EE	-	P	S+P	S+P *	S+P *
ES	S+(P)	S	S+(P)	S+(P)	S+P
FI	S	S	S	S	S
FR	S+(P)	S	N	S	S+(P)
GR	S	S	-	S	S
HR	-	-	-	NoE	NoE
HU	S *	S+P *	S+P	S+P	S+P
IL	NoE	NoE	NoE	NoE	NoE
IS	NoE	-	-	NoE	NoE
IE	-	NoE	-	-	-
IT	S	S	-	-	-
LU	S	S+(P)	-	-	-
LT	-	-	-	N	-
LV	-	-	S+(P) *	P	-
NL	S+(P)	S+(P)	(P)	(P)	(P)
NO	N	S+(P)	S	P *	P *
PL	NoE	NoE	NoE	NoE	NoE
PT	N	S	S	S	S
RO	-	-	S	S	-
RF	-	-	S+P	S+P	S+P
SE	S+P *	P *	S+P *	P *	(P) *
SI	S *	S+(P)	S	S	S
SK	-	S	S	S	S
TR	-	S	-	S+(P)	-
UA	-	NoE	NoE	NoE	NoE
UK	NoE	NoE	NoE	NoE	NoE

N = no issues (correspondence between missing value cells on sample and control data);
NoE = no 3-dimensional GAE table (education is in a separate table)
S = there are missing value cells on sample that do not exist in control data;
P = there are missing value cells on control data that do not exist in sample data;
() = missing values do not exceed 1%
***** = missing value on control data for at least 1 point higher than on sample

2.3.1 Missing values in sample cells (e.g. Belgium)[7]

To show how to deal with missing values in sample data we take the case of Belgium (R1) which has many missing values in ESS data, while having very complete control data (LFS), with no unknown values on gender, age or education (Table 8).

Table 8: Control data (LFS) for Belgium (R1)

Gender	Age \ Edu.	Low	Medium	High	Total
Male	15-34	6.020346	6.527969	3.267913	15.81623
	35-54	7.152322	5.879319	4.946907	17.97855
	55+	9.308003	3.022759	2.315014	14.64578
Female	15-34	4.950064	6.145643	4.286181	15.38189
	35-54	7.110552	5.70174	4.822227	17.63452
	55+	13.62885	3.131248	1.782943	18.54304
Total	Total	48.17014	30.40868	21.42118	100

On the other hand, ESS data for Belgium have missing values on all three control variables, gender, age and education (Table 9). The size of the crosstab on sample has more cells than on control data - sample cells that don't exist on the population data are shaded.

Table 9: Sample data (ESS) for Belgium (R1)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	0.421274	0.263296	0.105319	0.105319	0.895208
	15-34	0	0	0.157978	0.105319	0.263296
	35-54	0	0.052659	0	0.052659	0.105319
	55+	0	0.210637	0.052659	0	0.263296
Male	miss	0.157978	0.263296	0.684571	0.789889	1.895735
	15-34	0.157978	4.107425	7.582938	4.949974	16.79831
	35-54	0.263296	4.791996	7.740916	4.949974	17.74618
	55+	0.052659	7.003686	4.160084	3.054239	14.27067
Female	miss	0.052659	0.473934	0.526593	0.73723	1.790416
	15-34	0.210637	3.633491	6.213797	4.212744	14.27067
	35-54	0.105319	5.371248	6.371775	6.213797	18.06214
	55+	0.157978	8.320169	3.212217	1.948394	13.63876
Total	Total	1.579779	34.49184	36.80885	27.11954	100

[7] Not that the example of Belgium that we present was prepared before the consultation of national coordinators that advised using a specific procedure for Belgium (GR, AR, ER). The example shown assumes Belgium will use the same procedure as other countries (GAE, R).

To prepare control data for weighting, the LFS table is expanded to the size of the ESS table and the values in shaded cells (missing values) are copied to corresponding cells. As a result, the sum of values in the control table exceeds 100% and other cells are proportionally adjusted (multiplied with the rate of the decreased sum to the total sum) to preserve the total sum, while keeping the ratio between existing cells (Table 10). It is assumed that values miss at random (MAR), given of course the observed data.

Table 10: Control data table expanded with missing values from sample (Belgium, R1)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	0.421274	0.263296	0.105319	0.105319	0.895208
	15-34	0	0	0.157978	0.105319	0.263296
	35-54	0	0.052659	0	0.052659	0.105319
	55+	0	0.210637	0.052659	0	0.263296
Male	miss	0.157978	0.263296	0.684571	0.789889	1.895735
	15-34	0.157978	5.649424	6.125772	3.066572	14.99975
	35-54	0.263296	6.711657	5.517086	4.642121	17.13416
	55+	0.052659	8.734524	2.836523	2.172382	13.79609
Female	miss	0.052659	0.473934	0.526593	0.73723	1.790416
	15-34	0.210637	4.645084	5.767002	4.022103	14.64483
	35-54	0.105319	6.672461	5.350448	4.525123	16.65335
	55+	0.157978	12.78916	2.938328	1.673093	17.55856
Total	Total	1.579779	46.46613	30.06228	21.89181	100

Next, values were adjusted to sum 1899, the sample size (ESS) for Belgium (Table 11). This table was used as the population table when post-stratifying ESS data.

Table 11: Control data table values adjusted to sample size (Belgium, R1)

Gender	Age \ Edu.	EDUCmiss	Low	Medium	High	Total
missing	Miss	8	5	2	2	17
	15-34	0	0	3	2	5
	35-54	0	1	0	1	2
	55+	0	4	1	0	5
male	miss	3	5	13	15	36
	15-34	3	107.2826	116.3284	58.2342	284.8452
	35-54	5	127.4544	104.7695	88.15388	325.3777
	55+	1	165.8686	53.86556	41.25354	261.9877
female	miss	1	9	10	14	34
	15-34	4	88.21015	109.5154	76.37974	278.1053
	35-54	2	126.71	101.605	85.93208	316.2471
	55+	3	242.8661	55.79884	31.77204	333.437
Total	Total	30	882.3919	570.8827	415.7255	1899

2.3.2 Missing values in population cells (e.g. Sweden)

When missing values are found in population cells the situation is more complex. Most simply, missing at completely random assumption (MCAR) is taken which allows ignoring unknown values. On the other hand, for cases where unknown values exceed 1% of the population, an alternative is re-allocation of cells to known values.

We take the case of Sweden which has the highest total sum for missing values that appear in control data (Table 12) but not in sample data (Table 12). The problematic five cells are shaded. These units have known values of gender and age but education is unknown.

Table 12: Control data (LFS) for Sweden (R1)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	0	0	0	0	0
male	15-34	1.331925	3.450051	8.274535	2.513832	15.57034
	35-54	0.112274	3.338191	9.966818	3.612453	17.02974
	55+	4.268347	4.880743	5.075639	2.272438	16.49717
female	15-34	1.292119	3.112317	7.445192	3.117707	14.96733
	35-54	0.114394	2.300435	9.041333	4.987296	16.44346
	55+	6.79064	4.856692	5.159642	2.684988	19.49196
Total	Total	13.9097	21.93843	44.96316	19.18871	100

On the other hand, these five missing data cells are blank in sample data (Table 13).

Table 13: Sample data (ESS) for Sweden (R1)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	0.2001	0	0	0	0.2001
male	15-34	0	3,151576	10,25513	2,251126	15,65783
	35-54	0	3,751876	10,70535	4,252126	18,70935
	55+	0.050025	9,504752	3,801901	3,001501	16,35818
female	15-34	0	3,651826	7,603802	3,051526	14,30715
	35-54	0	2,451226	7,753877	6,053027	16,25813
	55+	0	11,70585	3,401701	3,401701	18,50925
Total	Total	0,250125	34,21711	43,52176	22,01101	100

In addition, there is one missing cell that appears only on sample data, like in Belgium – we treat it as shown in Section 2.3.1: we insert it into the control table and rescale other values so that it sums to 100% (Table 14).

Table 14: Control data table expanded with missing values from sample (Sweden, R1)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	0.2001	0	0	0	0.2001
male	15-34	1,32926	3,443148	8,257978	2,508801	15,53919
	35-54	0,112049	3,331511	9,946875	3,605224	16,99566
	55+	4,259806	4,870977	5,065482	2,267891	16,46416
female	15-34	1,289534	3,106089	7,430294	3,111468	14,93739
	35-54	0,114165	2,295831	9,023242	4,977316	16,41055
	55+	6,777052	4,846974	5,149317	2,679615	19,45296
Total	Total	14,08197	21,89453	44,87319	19,15032	100

Until this step the procedure is the same, regardless of the assumption we take, MAR or equal distribution of unknown values. The follow up depends on the choice:

1) MCAR assumption (ignoring missing values)

Table 15: Control data table with ignored missing values (Belgium, R1, assumption 1)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	0.221404	0	0	0	0.221404
male	15-34	0	3.809721	9.137161	2.7759	15.72278
	35-54	0	3.6862	11.00587	3.989054	18.68112
	55+	4.713325	5.389564	5.604777	2.509342	18.21701
female	15-34	0	3.436778	8.221358	3.44273	15.10087
	35-54	0	2.540256	9.983898	5.507225	18.03138
	55+	0	5.363005	5.697538	2.9649	14.02544
Total	Total	4.934729	24.22552	49.6506	21.18915	100

Missing cells in the control table that do not appear in sample data are deleted and other values are rescaled so that the total sum 100% and ratios are preserved (Table 15). Then, values were adjusted to 1999, the ESS sample size for Sweden in Round 1 (Table 16).

Table 16: Control data table values adjusted to sample size (Sweden, R1, assumption 1)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	4.4258588	0	0	0	4.4258588
	15-34	0	76.156332	182.65184	55.490245	314.29842
	35-54	0	73.68713	220.00725	79.74118	373.43556
	55+	94.219369	107.73738	112.0395	50.161738	364.15799
	15-34	0	68.701186	164.34494	68.820173	301.8663
	35-54	0	50.779724	199.57813	110.08942	360.44727
	55+	0	107.20648	113.89378	59.268345	280.3686
Total	Total	98.645228	484.26823	992.51544	423.5711	1999

2) MAR assumption (re-allocation of missing cells to known values)

Table 17: Control data table redistributed missing values (Sweden, R1, assumption 2)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	0.2001	0	0	0	0.2001
male	15-34	0	3.765235	9.030466	2.743486	15.53919
	35-54	0	3.353621	10.01289	3.629151	16.99566
	55+	4.259806	4.870977	5.065482	2.267891	16.46416
female	15-34	0	3.399572	8.132354	3.405459	14.93739
	35-54	0	2.311915	9.086454	5.012185	16.41055
	55+	0	7.438362	7.90235	4.112246	19.45296
Total	Total	4.459906	25.13968	49.22999	21.17042	100

Also according the second assumption missing values are deleted but the treatment differs in how they were re-allocated. Instead of proportionally rescaling all other values at the same time, we do it cell per cell. For instance, the 1.32926 male, 15-34 years old, unknown education are proportionally distributed among the known education cells (low, medium, high) in the same gender and age category. As a result, since we use the known information in "missing" cells (e.g. gender and age in this case), the structure better reflects the information that is known from sample data (Table 17). Finally, same as in the first method, values are adjusted to 1999, the ESS sample size for Sweden in Round 1 (Table 18).

Table 18: Control data table values adjusted to sample size (Sweden, R1, assumption 2)

Gender	Age \ Edu.	miss	Low	Medium	High	Total
missing	miss	4	0	0	0	4
Male	15-34	0	75.26705	180.519	54.84228	310.6283
	35-54	0	67.03889	200.1576	72.54672	339.7432
	55+	85.15353	97.37083	101.259	45.33515	329.1185
female	15-34	0	67.95743	162.5658	68.07513	298.5983
	35-54	0	46.21518	181.6382	100.1936	328.047
	55+	0	148.6929	157.968	82.20379	388.8646
Total	Total	89.15353	502.5422	984.1076	423.1967	1999

2.4 Weighting procedure in R

In most of the cases, we have control data for gender, age, education and region, including their interaction; we post-stratify the three-dimensional (interaction) GAE table (Gender, Age, and Education) and rake it with the Region variable. Although data would sometimes permit also a four-dimensional table post-stratification, we decided to have region in a separate table to avoid the problem of having empty cells, provide comparability and keep the weighting procedure smooth and simple.

After adjusting the population data to the missing values, we continue with the weighting procedure using the "Survey" package in R (Lumley 2010). Dweight is used as the initial weight. The weighting procedure is described in Figure 3 (example for Slovenia 2002).

However, this procedure was not possible in all countries, First, in some countries region was recoded to one category in NUTS2 (Cyprus, Estonia, Latvia, Lithuania) or original had only one category in ESS data, also NUTS2 (Iceland, Luxembourg). In this cases only post-stratification on gender, age and education is performed, no raking.

Second, we had to modify the procedure for Croatia, Norway, Israel, Russia and Ukraine where another control data source was used instead of LFS. In these cases the interaction is known only for gender and age, while education frequencies are given separately. Thus, a three-table raking is needed, where the first table is gender by age, the second education and the third region.

Figure 3: Example of weighting procedure in R (Slovenia 2002)

```
library(survey)

popt1 <- data.frame(demogr=c(31, 110, 111, 112, 113, 120, 121, 122, 123, 130, 131, 132, 133, 210,
211, 212, 213, 220, 221, 222, 223, 230, 231, 232, 233), Freq=c(0.0658327847267939,
0.164614599574886, 5.20854213015403, 11.1258481448688, 1.16582147070554,
0.318313567241112, 3.3685118867948, 12.2326393233638, 2.26714891049422,
0.24408667798703, 3.82987649818321, 6.94737149489314, 1.58198126984237,
0.414494913553804, 4.46580534020639, 9.49246449509759, 2.14137394161245,
0.363603074846074, 4.94545704062419, 9.83995886909441, 2.58801873749212,
0.13052146232268, 9.83139101607267, 6.16355923537798, 1.10276311486988))

popt2 <- data.frame(regionRx=c(01, 02), Freq=c(48.9164086687306, 51.0835913312694))

ess_data=read.csv2(file="SI2002.csv")

dclus1 <- svydesign (ids=~1,data=ess_data, weights=~dweight)

raking <- rake (design=dclus1, sample.margins=list (~demogr, ~regionRx),
population.margins=list(popt1, popt2))

weights(raking)

write.table(weights(raking), "wSI2002.csv")
```

2.5 Weight trimming

In some countries weights can have large variation. It is of course not a very good idea to give a unit too much weight in the dataset, due to the potential increase in variance, but also in bias. For these reasons trimming procedures are used, so that the weight has no values above a certain number. After cutting all the higher values the weights is rescaled so that the average 1 is preserved.

In ESS weighting we cut weights at 2 and 4 which all decreased the variance inflation factor (VIF), with very little differences among the approaches. However, this also shrinks the potential improvement and removal of non-response bias, so it has to be done very carefully. The decision on the optimal cut-off level is largely arbitrary, although it is true that large survey organization (e.g. statistical offices) are usually very conservative and typically cut the weights at 2, while some (e.g. Slovenian Statistical office) also cut at 5.

Of course, if we cut exceedingly we also risk moving too far from the target controls margins. Thus, a careful analysis of the effect of weight trimming on control data is needed to estimate if the data should be re-weighted.

For now we only perform trimming for analytical purposes and we will not propose any restrictions, as we believe that the point estimates are the closest to the true value.

Nevertheless, before the data files are given to public users this solution might be reconsidered, based on comparison of (a) no trimming, (b) trimming at 2 (c), trimming at 4, which is also the proposed approach.

The effects of trimming of the weight on variance, bias – as well as on the discrepancies in the matching of the control variables - are in the process of calculation, together with the analysis of the effects of the weights. They will be finished by the end of September 2012 and also presented accordingly. Nevertheless, first results show that the differences are relatively minor.

3 Open issues

Post-meeting comment: *Most of the raised issues were resolved at the ESS Post-stratification Weights meeting in London on October 2nd 2012. Please see Appendix E.*

Summarizing the above elaboration, we can structure the dilemmas in conceptual and specific. In general, with respect to all issues, we basically propose to keep with the solutions that we have used or proposed so far by default.

a) General issue: Are the four control variables sufficient and is thus the existing weighting scheme final? Sometimes in some countries also other variables could be used in weighting, e.g. nationality, household size, degree of urbanization, occupational status, etc. Of course, we could go even one step further and depart from the standardized set of ESS weighting variables and enter into country specific weighting, which is now, of course, in a certain (unexplored) relation with the proposed ESS weights.

b) Education variable:

Education – conceptual issues: There are problems with the ISCED category 3C, i.e. lower vocational education, which we already elaborated in details in previous sections. There are five possible solutions:

1. We could continue with existing strategy with additional improvements coming from better education data in R5 onward. This is also our default proposal, supported also by ESS educational experts.
2. We could continue with the existing strategy but also acquire more detailed LFS data from national offices in each country, in order clarify more the issue with 3C. Here, our estimate is – supported also by ESS education experts - that there would be a lot of efforts, while the added value would be very questionable;
3. Another possible approach – mostly endeavored by ESS education experts - would be to treat 3C category separately and have a four-category education control variable in weighting. In fact, from a sociological perspective, 3C is very often a completely different category than the rest of the Medium level (3A and 3B), which can proceed with the education at ISCED level 5. At the same time, very

often, this category is also very different from the Low level (ISCED 1, 2). Moreover, its role also varies across countries, e.g. in Germany it is very similar to ISCED 3B, while in Slovenia it is closer to ISCED 2 (Low), than to 3B (Medium). Thus, the separation of 3C would be beneficial from many substantive points. Of course, as we mentioned above, extra efforts would be required at the ESS and LFS data side, but also from the weighting aspect.

4. Another option would be to put entire ISCED 3C into Lower level, as socio-economic and attitudinal characteristics of this segment are much closer to other Lower level segment than to Medium level (i.e. 3A and 3B). In fact, this is actually the categorization often used in marketing research. However, a problem is that in some countries the 3C is very similar to 3B (e.g. Germany).
5. We might also depart from LFS and apply other sources, i.e. direct contact of the countries (e.g. modified Census data etc.).

Education – specific problems:

- The situation with the post-secondary, non-tertiary ISCED 4 category is not always clear enough. Different countries do not follow the official ISCED rules which causes them having no ISCED 4 category at all (e.g. Slovenia) or including in it something that is not really ISCED 4. This could affect our weighting accuracy. However, we consider that it suffices to check that the ESS and LFS distributions are similar enough. There were mismatches in certain countries which we tried to resolve. We discuss this in the next paragraph. Apart from this we decided to not go so much in detail – if there is a mistake it is the same mistake in ESS and LFS data and should not affect the weighting. However, this decision may be re-considered. This is an important dilemma.
- Using an **alternative control data source for education** in Israel, Russia and Ukraine (all rounds). As now we use census data for Ukraine and ESS Appendix (data provided by national coordinators) for Israel and Russia. Are there better alternatives?
- Using **control data for education from neighboring rounds** for Switzerland (R5) and Turkey (R2). Does this intervention have any effect on data quality?
- Issues with **adjusted education recoding of ESS data to achieve better match with control data** (ISCED categories) in United Kingdom (all rounds), Poland

(R5), Lithuania (R4), Israel (R1 and R4), Ukraine (R2, R3, R4and R5) and Russian Federation (R3, R4 and R5). In these countries the ESS recoding of education data intentionally deviates from official ISCED mappings used in control data (LFS). For weighting purposes the ESS recoding was changed to achieve better correspondence between sample and control data. In this way intra-country comparability would not be decreased. However, this is only a back-recoding for weighting purpose, while ESS stays as it is.

c) **Age variable – specific issues:** We have incomplete data for the age variable for Iceland and Norway – in LFS there is no data for age category 75+. Is the ESS appendix an appropriate solution or should we look deeper into it?

d) **Region**

Region conceptual issues: Is the existing prevailing weighting with region at NUTS 2 level sufficient or should we try even harder to attain more detailed data for all countries? (E.g. attain NUTS 3 or national specific regional data from other sources)? How to deal with comparability across rounds, as regional levels often change?

Region – specific issues are related to the change of regional classification for certain countries. For instance, in Round 5, Belgium and France have a higher number of categories than in previous rounds (e.g. Belgium, France). Similarly, Greece and Switzerland have a lower number of categories in one of the rounds (CH in Round 1, GR in R4), while other rounds have a higher number of categories. The dilemma is: Should we recode the variable to a lower number of categories to keep comparability with previous rounds, even if more detailed data is available, or should we exploit this new information but reduce comparability with rounds that need this recoding? In our opinion, the first solution is preferable.

e) **Missing values**

Missing values – conceptual: is the presented approach to handle the missing basically acceptable? Or imputations instead of weighting may be considered?

Missing values – specific issues: we assume missing at random (MAR) only when the sum of missing values exceeds 1%, as much more work is need in such case. The dilemma is the threshold 1% which worked well in previous exercises – no differences were found if we prepare two post-stratification control tables and compute the weights separately for both assumptions. The alternative would be 0.5%, 0.1% or 0.0%, what would of course require extra work.

f) **Weighting method** - is the post-stratification approach sufficient or should we use even more complex weighting procedures (calibration methods)? As mentioned above – with post-stratification, should we strive to also include region into the post-stratification table, as a fourth dimension, whenever possible (instead of raking)? So far, this was not the case due to potentially small cells and standardization of approach.

g) **Weight trimming** is related to the decision of the optimal cut-off level. In previous weighting exercises we evaluated the cut at 2 and 4 (which is the default approach) but the differences were not substantive. Initially, forthcoming preliminary weights will be run even without trimming. Then, due to several changes in ESS data and in procedures, this decision will be re-considered by analyzing increases in variance and reductions in bias to find an optimal solution. Nevertheless, our initial proposal is to trim at 4, which is slightly less conservative compared to usual practice in many statistical offices, but removes more bias. Another issue is what to do after the trimming (besides rescaling) – should there be some further fine tuning, smoothing or asymptotic adjustments of the weights? So far, this is not foreseen, as we estimate that these effects would be rather negligible.

4 Sources

ESS Appendix A1. EDUCATION, ESS5 - 2010 ed. e02_1. http://ess.nsd.uib.no/ess/round5/

ESS Appendix A4, POPULATION STATISTICS, ESS5-2010 ed. 2.0. http://ess.nsd.uib.no/ess/round5/

European Social Survey. Education upgrade. http://ess.nsd.uib.no/ess/doc/Education_Upgrade_ESS1-4.pdf

Eurostat NUTS. Nomenclature of territorial units for statistics. Available at: http://epp.eurostat.ec.europa.eu/portal/page/portal/nuts_nomenclature/introduction

Ganninger, Matthias. 2007. Weigthing in the ESS Cumulative Dataset. Available at:

http://ess.nsd.uib.no/ess/doc/Weighting_ESS_cumulative_data.pdf

Labour Force Survey. Available at: http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/

Labour Force Survey: Levels of education and training in the EU Labour Force Survey. http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/ISCED_EN.htm

Labour Force Survey: Statistical classifications used in the EU Labour Force Survey. http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/lfs_main/lfs/lfs_statistical_classifications.htm

Lumley, Thomas. 2012. Survey: analysis of complex survey samples. R package version 3.28-2. Available at: http://cran.fhcrc.org/web/packages/survey/index.html

State Statistics Committee of Ukraine. All-Ukrainian population Census 2001. Available at: http://2001.ukrcensus.gov.ua/eng/results/general/estimate/

OECD. 1999. Classifying Educational Programmes Manual for ISCED-97 Implementation

in OECD Countries. http://www.oecd.org/education/highereducationandadultlearning/1962350.pdf

Rydland, Lars Tore and Ostensen, Ase Gilje. 2007. Contextual Data for the European Social Survey. An Overview and Assessment of Extant Sources. Norwegian Social Science Data Services. Available at: http://ess.nsd.uib.no/ess/doc/ESS_context_sources.pdf

UNESCO. ISCED 1997: http://www.uis.unesco.org/education/ISCEDmappings/Pages/default.aspx

Vehovar, Vasja. 2006. Weighting in ESS-Round 1. Internal materials. http://mi.ris.org/uploadi/editor/1166602083ESS_weights_clanek_R1.doc

Vehovar, Vasja and Tina Zupanič. 2007. Weightin in ESS-Round 1 and 2. Internal document. http://mi.ris.org/uploadi/editor/1172233696ESS_weights_paper_R2_final.zip

Vehovar, Vasja. 2007. Non-response bias in the European social survey. V: Geert Loosveldt, Marc Swyngedooq and Bart Cambré, Bart (eds.): Measuring meaningful data in social research, Acco, pp.35-356.

Vehovar, Vasja, Katja Javoršek, Meta Makovec in Ana Slavec. 2008. Population weighting in the ESS. Round 1- 3. Internal materials. http://mi.ris.org/uploadi/editor/1252231942WeightingESS2009c.rar

5 Literature

Bethlehem, Jelke and Silvia Biffignandi. 2012. Handbook of Web Surveys. Wiley and Sons Inc., Hoboken, New Jersey.

Billiet, Jaak, Akim Koch and Michel Philippens. 2007. Understanding and imrpoving response rates. In: Roger Jowell, Caroline Roberts, Rory Fitzgerald and Gillian Eva (eds.): Measuring attitudes cross-nationally: lessons from the European Social Survey, Sage Publications, pp.113-137.

Cobben F. 2009. Re-Approaching Nonrespondents. Chapter in Nonresponse in sample surveys : methods for analysis and adjustment. Dissertation. Available at: http://dare.uva.nl/document/146561

Dever, Jill A., Ann Rafferty and Richard Valliant. 2008. Internet Surveys: Can Statistical Adjustments Eliminate Coverage Bias? Survey Research Methods 2 (2), 47-62. Available at: https://ojs.ub.uni-konstanz.de/srm/article/viewFile/128/1982

Gallup Europe. 2010. 5th European Working Conditions Survey. Weighting Report. Available at: http://www.eurofound.europa.eu/surveys/ewcs/2010/documents/weighting.pdf

Gallup Europe. 2010. 5th European Working Conditions Survey. Overview Report. Available at: http://www.eurofound.europa.eu/pubdocs/2011/82/en/1/EF1182EN.pdf

Graham Kalton and Ismael Flores-Cervantes. 2003. Weighting Methods. Journal of Official Statistics, 19(2), 81–97. Available at: http://www.jos.nu/Articles/abstract.asp?article=192081

Groves, M. Robert. 2006. Nonresponse Rates and Nonresponse Bias in Household Surveys. Public Opinion Quarterly 70 (5), pp. 646–675.

Kish, Leslie. 1965. Survey sampling. Wiley.

Leslie Kish. 1992. Weighting for Unequal Probabilities. Journal of Official Statistics, 8(2), 183–200. Available at: http://www.jos.nu/Articles/abstract.asp?article=82183

Lumley, Thomas. 2010. Complex Surveys: A Guide to Analysis Using R. Wiley.

Meuleman, Bart and Jaak Billiet. 2004. Corrections for nonresponse in the ESS round 1: the use of weighting factors.

Schouten, Barry. 2006. The Selection of Strata in Nonresponse Adjustment . 10th International Blaise Users Conference. Available at: http://www.blaiseusers.org/2006/Papers/425.pdf

Schneider, L. Silke. 2009. Confusing credentials: The Cross-nationally Comparable measurement of educational attainment. Submitted for the degree of Doctor of Philosophy (DPhil). Available at: http://mi.ris.org/uploadi/editor/1395683827ThesisSilkeSchneider.pdf

Stoop, Ineke, Jaak Billiet, Achim Koch and Rory Fitzgerald. 2010. Improving Survey Response: Lessons Learned from the European Social Survey. Wiley.

Appendix A: Structure of the reports and documentation

I. Methodology development and data preparation

1. Introduction and Description of Work
2. Standardized Procedure
2.1 Control Data Sources
2.2 Control Variables
2.3 Strategy for Dealing with Missing Data
2.4 Weighting Procedure in R
2.5 Weight Trimming
3 Open Issues

II. Final Methodological Report

1. Introduction
2. Standardized Weighting Procedure
2.1 Control Data Sources
2.2 Control Variables
2.3 Strategy for Dealing with Missing Data
2.4 Weighting Procedure in R
2.5 Weight Trimming
3 Weighting ESS data
3.1 Round 1
3.2 Round 2
3.3 Round 3
3.4. Round 4
3.5. Round 5
4. Results
5. Summary

III. Analysis of PS Weighting Effects

1. Introduction
2. Nonresponse within ESS
3. Methodology
4. Weighting for nonresponse
4.1 Control variables
4.2 Computing weights
5. Effect on nonresponse bias
5.1 Relative bias
5.2. Standardized bias
5.2. Effect on VIF
6. Conclusions

IV. Public usage and notes for users (outline)

1. Data preparation
1.1 Description of ESS data and design weights
1.2. Description of control data (LFS)
1.2. Recoding of gender, age and education
1.3. Recoding of region
1.4. Dealing with missing data

2. Weighting implementation
2.1 Post-stratification weighting and raking in R
2.2 Weight trimming and scaling
2.3 Cross-national selection probability

3. Weighting evaluation
3.1 Analysis of weights
3.2 Analysis and evaluation of weighted data

4. Glossary

Appendix B: Slovenian coding of education in LFS 2010

	A	B	C	D	E	F	G	H
1	COUNTRY	SI						
2	YEAR	(All)						
3								
4	**Vsota od VALUE**	**Column Labels**						
5	**Row Labels**	**1. Low**	**2. Medium**	**3. High**	**No answer**	**Grand Total**		
6	0	57.30925				57.30925		
7	11	356.5752125				356.5752125		
8	21	2009.26847				2009.26847		
9	31		2139.984793			2139.984793		
10	32		2683.785415			2683.785415		
11	51			660.041455		660.041455		
12	52			598.57691		598.57691		
13	60			94.1532275		94.1532275		
14	NoAnswer				28.51691	28.51691		
15	**Grand Total**	**2423.152933**	**4823.770208**	**1352.771593**	**28.51691**	**8628.211643**		
16								
17								
18								
19	ISCED classification							
20	http://circa.europa.eu/irc/dsis/employment/info/data/eu_lfs/LFS_MAIN/Related_documents/ISCED_EN.htm							
21	0	No formal education or below ISCED 1						
22	11	ISCED 1						
23	21	ISCED 2						
24	22	ISCED 3c (shorter than 3 years)						
25	30	ISCED 3 (without distinction a, b or c possible, 3 y+)						
26	31	ISCED 3c (3 years and more)						
27	32	ISCED 3 a,b						
28	36	ISCED 3 or 4 (without distinction a, b or c possible)						
29	41	ISCED 4a,b						
30	42	ISCED 4c						
31	43	ISCED 4 (without distinction a, b or c possible)						
32	51	ISCED 5b						
33	52	ISCED 5a						
34	60	ISCED 6						
35								
36								

http://mi.ris.org/uploadi/editor/DnD1346849902ISCED97manual.pdf

http://mi.ris.org/uploadi/editor/DnD1347923394LFSISCED.xlsx

Appendix C: Task 3 in WP12 of Description of Work

Final version from Dow[8]

Wp12, Task 3: Post-stratification (UL, SCP)

A further task is the calculation of post-stratification weights that can correct for unequal representation of different socio-economic groups among the respondents. Earlier studies on post-stratification in the ESS showed that nonresponse bias (as measured by the size of the post-stratification weights) differed substantially across countries, and pointed to the importance of relevant and reliable post-stratification variables. However there remain a number of obstacles prior to deriving these weights, such as difficulties in using education consistently across countries. The objective of this task is to resolve these outstanding issues and provide improved PS weights for all rounds for all countries. The work will involve:

- making an inventory and assessment of variables available and their suitability for PS
- performing and publishing the post-stratification for round 5 (deliverable12.6)
- establishing a standardized procedure for harmonized post-stratification strategy, which should handle age, gender and education
- elaborating a strategy for dealing with missing data in sample and/or control dataset
- performing and publishing post-stratification weighting for all past rounds when a standardized procedure has been developed and providing an explanatory methodological note about using PS weights for public users of ESS data (Del 12.8)

[8] http://mi.ris.org/uploadi/editor/1277884892Paper3ESSDACEDoW-toEC.doc

Initially proposed Task 4: Post-stratification (UL, SCP)

A further task is the calculation of post-stratification weights that can correct for unequal representation of different socio-economic groups among the respondents. Earlier studies on post-stratification in the ESS (Vehovar 2007) showed that nonresponse bias (as measured by the size of the post-stratification weights) differed substantially across countries, and pointed to the importance of relevant and reliable post-stratification variables. However there remain a number of obstacles prior to deriving these weights, such as difficulties in using education consistently across countries. The objective of this task is to resolve these outstanding issues and provide improved PS weights for all rounds for all countries. The work will involve:

• making an inventory and assessment of variables available and their suitability for PS
• making an inventory of national weighting procedures in instances where national teams provide country specific weights
• performing and publishing the post-stratification for every new round (deliverables 12.7 & 12.15)
• establishing a standardized procedure for harmonized post-stratification strategy, which should handle age, gender and education
• elaborating a strategy for dealing with missing data in sample and/or control dataset
• performing and publishing post-stratification weighting for all past rounds when a standardized procedure has been developed and providing an explanatory methodological note about using PS weights for public users of ESS data (deliverable 12.10)

This is a continuation of a small task in the ESSi grant which was sufficient for exploratory work but never facilitated the actual production of weights for each round of the ESS. In this proposal we request funds to further develop the PS strategy and to apply this to former and future rounds.

Appendix D: Minutes from ESRA meeting

What: Minutes from the meeting, July 18th, at 2011ESRA conference

Who: Vasja Vehovar, Ineke Stoop, Silke Schneider, Ana Slavec, Hilde Orten, Eric Harrison, Hideko Matsuo, Jaak Billiet

When: ESRA venue, 18th July, 2011, 12.30-13.30

A. Participants first discussed EDUCATION. They agree:

1. It is important that ESS provides formal monitoring of the education variable for each ESS wave after this exercise is over.
2. It is equally important that the responsible person (currently this is **Silke**) is in regular contact with corresponding experts in OECD but also in Eurostat (which is a major source for control data on education). For this reason **Vasja** already forwarded the Eurostat contact, Mr. Bauer, and also the recent LFS data to **Silke**.
3. It is estimated (**Silke**) that by October 2011 the education problem documentation for round 5 will be ready. This also includes the control data for countries, where LFS is not a good source (UK, Germany, and Norway). After that, **Ana and Silke** will be in contact and will verify the education data for round 5 (in November).
4. **Silke** estimated that previous rounds (R1, R2, R3 and R4) were also successfully corrected, except for few cases, which will be corrected with some delay, i.e. within a year (HU, SE ...) the latest, and also for few other cases where LFS is not a proper source (UK, NO ...). **Ana and Silke** will clarify and confirm the data (based on Slovenian documentation bellow) for each country and each round. After that the Slovenian team will finalize calculation of corresponding weights for those countries.

B. REGION

1. Vasja will re-send the Slovenian study (enclosed at the bottom of this message), which also discusses the region variable, to the Norwegian team (Christine, Hilde) as well as to the German team (Siegfrid, Sabine), so that they check the availability of the region variable at the units level. The Norwegian team will report on that within a month. If a problems exsist, the region variable will not be used for previous rounds.
2. Hilde assured that the new archiving system for linking units with regions is expected to be operational by the end of 2011. After that it will be checked in order to outline the standardized usage for weighting purposes.
3. Vasja will prepare the exact weighing scheme for ESS at the methodology meeting in Manheim in October 2011.
4. The final strategy for ESS population weighting will be done in 2011. It needs to be approved and verified by ESS bodies (CCT, Methods Group ...). After that weighting will be finalized as foreseen in the DACE proposal DoW.

Link:

http://mi.ris.org/uploadi/editor/DnD1311693176WeightingESS2009c.doc (ESSi study on weighting)

Appendix E: Minutes from London meeting

Content: **Minutes ESS Post-stratification Weights meeting – Draft 2**

Location and time: **City University, October 2nd 2012, 9:30-12:30**

Attendees: Vasja Vehovar, Ana Slavec, Ineke Stoop, Matthias Ganninger, Verena Halbherr, and Ana Villar.

A. General information:

1. *National weighting procedures*. It is proposed that it may be necessary to consult with countries on what they do for weighting: which variables they use and how they compute weights[9].
2. *Quality of population benchmarks.* We had a brief initial discussion of LFS as source of population values. It is remarked that the survey has large sample sizes (about 20,000 respondents per country per month), and that quality has been increasing. The fact that LFS data are weighted is considered reassurance that population estimates are appropriate for age and gender (as stated in the document circulated before the meeting).[10]
3. *Non-LFS countries*. In countries for which no LFS data are available, data provided by NC's for Appendix 4 of the Documentation Report are used as population benchmarks. At the meeting it was suggested that users' documentation should include a citation of this report, in addition to information about the source of the data. Also, it was discussed that someone at CST should try to monitor the quality of the data provided by the NC's regarding population benchmarks. (Data missing from both LFS and Appendix 4 were obtained if possible from censuses.)
4. In some instances there is no cross-tabulated data for age*gender*education. Two possible solutions: 1) rake to marginals of each variable; 2) use cross tabulations from other years to estimate what they would have been for the year that is missing the information, maintaining the marginals of the new year. The second option seems to be the one that will be implemented.

B. Recoding education to match categories of LFS and ESS

1. Numerous problems arise from making the variables of education in the LFS and ESS match. Different countries have different problems. Some ISCED classifications are difficult to assign to one of the three levels used for weighting.
2. In addition, decisions and considerations applicable to one year one country might not apply to other years and countries. This is due to a) changes in educational system; b) changes in the ESS variables across time; c) changes in the LFS variable across time.

[9] Quality report: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-11-020/EN/KS-RA-11-020-EN.PDF. National reports: http://epp.eurostat.ec.europa.eu/cache/ITY_OFFPUB/KS-RA-12-013/EN/KS-RA-12-013-EN.PDF Questionnaire: http://www.esds.ac.uk/doc/7059%5Cmrdoc%5Cpdf%5Clfs_user_guide_vol2_questionnaire2011.pdf Generic information: http://epp.eurostat.ec.europa.eu/portal/page/portal/employment_unemployment_lfs/methodology/organisation

[10] However, in 2010 only Denmark adjusted weights to match the distribution of educational level in the population.

3. *Silke Schneider's* help has proven very useful and she is being consulted every step of the way when they find problems recoding.
4. It is decided that all decisions and changes to educational variables (coding scheme, and so on) should continue to be documented.
5. LFS and ESS differences were especially large for UK, partly due to adjustments done in the ESS to make UK data more comparable to other countries. When inspecting this problem at the meeting, it was noticed that the UK LFS variable of education had a large percentage of missing values (15% and up). This raised serious concerns about the quality of LFS in the UK, and whether it should be used as a benchmark.
   Post-meeting action point: At the CST meeting it was decided that Ana Villar and other group members would consult with national experts about sources of education data in those countries where the LFS educational variable shows high levels of missingness (see DACE 9th meeting minutes).
6. It was discussed that population benchmarks with missing values higher than 1% might not be appropriate sources.
7. The percentage of missing values for the education variable in control (LFS) was checked, and about ten countries had levels of missing in the education variable higher than 5%. On the other hand, there is only one country in one round (Ireland, 3rd round) with at most 5% of missings for education in ESS. It was discussed whether we could assume that they are MCAR or MAR. The default approach so far in this project has been to use MAR.
8. Different possible solutions for treating high item nonresponse rates (as in UK for LFS, where it was 29% in 1st round and it still 15% in 5th round) were identified and discussed: multiple imputation, single imputation (hot deck, regression, nearest neighbor), use available published imputation if existing, using alternative sources, not use education (or the problematic variable) to weight in that country. It was decided that Vasja Vehovar would consult with substantive experts (Silke Schneider) and country experts (Peter Lynn) on this to try to identify why LFS data in the UK have have such high level of missingness and how the issue could be best solved. In addition it is decided that an analysis of weight inflation will be conducted (see weight evaluation below).
9. Need to make a decision of what to do if and when LFS data are deemed not ok.
10. Regarding the number of categories that should be used to weight for education, it is decided that 3 categories is - given resources and time constraints - the optimal approach at this stage, which is also in line with accepted practice in previous weighting activities, but that for computation of future round weights this decision could be revisited.

C. Recoding region to match categories of LFS and ESS:

1. Region is considered a proxy for level of urbanicity. It is not clear whether this assumption is valid for all countries.
2. *Optimal level of NUTS for weighting.* It was discussed what level should be used to weight to region estimates, whether NUTS2 or NUTS3 is more appropriate.
   - NUTS2 represents too large regions in some countries (large countries) as compared to NUTS2 regions in other smaller countries.
   - At the same time, NUTS2 level in many countries includes more than 10 regions. The larger the number of regions, the more likely that the number of ESS sampled individuals in that particular region will be smaller, threatening robustness of the weights (see discussion on *thresholds*).

- Using NUTS3 would not be possible in some countries, because data are not available in ESS, LFS, or both.
- *AV: I have checked the LFS website after our meeting, and an additional reason not to use NUTS3 is that LFS data in 12 countries are weighted to match population the distribution of the population across NUTS2 regions. Therefore, that is the level at which our benchmark data are most accurate.*

3. *Number of regions*: The discussion of whether to use NUTS2 or NUTS3 led to a discussion of how many categories where optimal for weighting. The rationale is that more categories help create, in principle, weights that result in less biased estimates, but they can also increase weights variance. Therefore, there is a trade-off between bias and variance. So it is proposed to come up with a rule-of-thumb criterion for the minimum sample cell size to use for weighting.
4. One option - to be rechecked with external experts (Vasja Vehovar) is as follows: A 30 cases and 5% minimum criterion is proposed where both conditions need to be met (PIAC follows this approach, ANES follows a 5% minimum approach). When sample cells are smaller than 30 or represent less than 5% of the sample, they will need to be merged with the smallest neighboring region. Exceptionally, if there is one cell only below 5%, and that cell is above 4%, this would not need to be recoded/merged.
5. It is decided to continue using NUTS2 level regions to compute the weights, because NUTS3 would lead to many cells that do not meet the criterion.
6. *Within country changes of NUTS level*. It was discussed what to do when ESS or LFS data are at one NUTS level for one round, but at a more detailed level for another. The suggestion was to use the best possible option each year (NUTS2 as per current specification).
7. In the future, however, a more detailed region variable is to be explored and considered, so to take the advantage of this important variable, which is typically used in national weighting approaches.
8. Although out of the scope of this task, it would be interesting to get a better sense of what NUTS represent in each country, how they have been created, and which what types of variables they correlate.

D. Trimming:

1. According to the analyses presented on the report, it is considered that trimming weights at 4 is an appropriate approach for the ESS post-stratification weights. In addition, the design weights are also trimmed at 4, as well as weights from other high quality surveys.
2. In case of serious increase of variance, further refinements of the weights should be checked (eg. Anesrake), if they improve the situation.

E. Weight evaluation:

1. Sensitivity analyses: test in three (or five) problematic countries what the effect is of using different approaches on weights and estimates.
2. Inflation of weights analysis: to test whether missing-ness affects weights, compare what weights would be if all those missing would have "belonged" to the smallest cell to what weights would be if all the missing would have been in the largest cell. If the difference in inflation of weights is less than 1%, we could consider the issue ignorable.
3. For Slovenia the test weighting will be also done for 4-level education and NUTS3 region to check the impact on estimates.

F. Weights release and documentation:

1. It was said that among the additional variables to be included in the dataset where weights are released are the variables that were used to compute the weights (the recoded education and region variables, and even gender and age, including any imputed values where imputation might have occurred), the original variables from which these were recoded, and a code sheet explaining how recoding was performed.
2. Discussion about whether to include weights that are independent from each other as they are provided now (design alone, population alone) or combine them. There are divided opinions. However, the post-stratification weights *will* include the design weights. So we need to make very clear how each weight was computed and should be used. We could have dweight and pweight as until now, and add dpweight, postweight, ppostweight.
3. Similarly, it was discussed whether the ESS website should activate the weights by default or if it should be the user's decision. This decision is to be revisited at the next CST meeting.

G. Next steps:

1. *Weighting experts*. It is decided that Vasja Vehovar will contact weighting experts about (a) trimming (see D1), (b) minimum cell/margin size for raking/post-ratification (see C4), and (c) missing level at population controls (see B8). Vasja Vehovar will first circulate the questions among those at the meeting, so that the participants will ask various experts for their practice.
2. *Timeline.* Vasja Vehovar considers that things are progressing at the right pace to meet the timeline provided in the report circulated at the meeting.
    a. The initial weights for round 5 are now calculated for all countries and are available together with corresponding documentation (HERE URL), however they will be further refined and processed after the remaining methodological decisions are taken,
    b. Except for the three open issues above (G1) the participants agreed that the solutions in the "**Post-stratification weighting of the ESS** - Methodology development and data preparation, DACE project, activity of WP12, September 2012" (available here) – together with corresponding analyses of weights and calculation of biases, and with further clarifications within this minutes - are accepted. The only remaining methodological clarifications should be done in October, while the remaining tasks from D and E will be done in November.
3. *CST:* evaluate ESS capabilities to deal with checking the quality of the population benchmarks provided by countries with the technical report.
4. Matthias Ganninger*:* Proposal for the variables to be included into final (public) data set, proposal for final documentation about weight (to replace the existing one, which is currently here http://ess.nsd.uib.no/ess/doc/weighting.pdf)

H. Resources

Vasja Vehovar also explicated that this WP12 task basically assumes the weighting activities to be done according to the strategy outlined in DoW (also available at the end of the document here[11]) and it in principle already excludes any further iterations that might be needed for primary data quality control checks.

We were thus assuming here that we would obtain readily coded ESS control variables, which were to be weighted towards control variables (where we decided to use LFS).

Initially, the DoW has even foreseen weighting only for age, gender and education, which was then agreed to be - in addition - extended by region at the meeting of the WP12 group in 2011 (at ESRA conference meeting).

However, within the process of reduction of resources - during the DACE application in the revised DoW - the subtask of contacting the national sources and their weighting practice was entirely omitted. Consequently, to go and to contact national coordinators and the corresponding weighting experts, to make the inventory of their approaches, to compare their approach with ESS weighting and to further clarify potential issues related to primary data quality (as it is now already the case with UK) – this is now all out of the task 3 in WP12 of the current DoW (although was foreseen in initial DoW, also included in the document).

This is particularly problematic, because after the weights will be distributed to the users, we can expect almost for sure a continuous flow of comments from national users (particularly those who know some specifics for their country), which might sometimes also need further clarifications and even weighting re-runs.

In addition, there are already now numerous discrepancies between ESS and LFS data which are obviously calling for profound national clarification already at this point (e.g. region in Croatia, Slovakia, Turkey, Ukraine …. education in UK, Sweden, Denmark, Estonia, Hungary, Latvia …), so it is very likely than one day some users will raise these issues.

Meeting concluded at 12.30.

[11] http://mi.ris.org/uploadi/editor/DnD1349079494ess_weight_OCTx.pdf

Post meeting comments (by Vasja Vehovar)

In fact, after years in this weighting exercise, I see that weighting itself is not a problem, but the data clarification is a problem. So perhaps the right strategy would have been - at very first - to obtain a competent national consent about the variables and the margins, and only after that to proceed with weighting. Now we are trapped again into the situation that we do not know at all what national users will say.

In any case we are here at UL somehow reluctant to deal with further comments and requests coming in the future from the countries, or, even with requests to systematically check the weighting with countries beforehand.

This could be done only with increased resources, as initially foreseen for this task.

Post meeting examination of LFS (by Ana Villar)

Summary on LFS methodology (based on a methodological reports of the 2010 survey). This section intends to complement the information provided in the report, but some of the information was already presented there.

1. As stated in the report, sample size is large, sampling design is probabilistic.
2. Participation is mandatory in 14 of the countries.
3. Sampling frame is either the census and/or a population register. Exceptions: Luxembourg and Switzerland use a telephone list (Switzerland adds a register of foreigners), UK an address list (with telephone list for northern Scotland), the Netherlands and Turkey complement their register with an address list.
4. PROXY INTERVIEWS ARE ALLOWED. The range of percentage of proxy responses ranges from 2% in Switzerland (Germany 3.6% and Israel 2.6%) to 57.2 in Slovenia, with an average of 31%. Please see UK report on proxy responses below.
5. Response rates tend to be high:
    - > 90%: Romania (voluntary), Slovakia, Austria, Germany, and Cyprus.
    - 81%-90%: 15 countries (8 compulsory, 7 voluntary).
    - 71%-80%: 8 countries, 5 voluntary, 3 compulsory.
    - 61%-70%: Estonia and Latvia (voluntary).
    - 51%-60%: Denmark and UK (voluntary). It looks like UK response rate at wave 1 might be around 71%[12].
    - 31%: Luxembourg, voluntary.
6. Usually all members of the household 15+ are interviewed (individuals are the final sample unit in Finland, Denmark, Switzerland, and Sweden)
7. Weights take into account probability of selection, and distribution of the population with regard to age (5-year age groups, except in 6 countries), gender (except Croatia), and region (NUTS3 for 21 countries, NUTS2 level for all others). **18 countries gross the sample to the total population including people living in institutional households.**
8. Data collection methods:
    - 16 countries collect the first wave F2F and the rest by telephone.

[12] http://www.ons.gov.uk/ons/guide-method/method-quality/specific/labour-market/labour-market-statistics/volume-1---2011.pdf

- 8 countries collect only F2F (Bulgaria, Estonia, Ireland, Portugal, Romania, Croatia, Macedonia, Turkey).
- 6 countries collect only by telephone (Luxembourg, Finland, Sweden, Iceland, Norway, and Switzerland).
- Germany F2F, some refusals telephone/mail.
- Belgium: F2F, households of retired people can be conducted by telephone.
- Denmark: Telephone for individuals, web for other members of the household.
- Italy: Telephone if number available; F2F for no known number and foreigners.
- Seven countries use paper questionnaires (mixed of PAPI and mail???) (Bulgaria, Greece, Hungary, Romania, Slovakia, Croatia, Macedonia).

9. Length of interview: 31 minutes on average for first wave, 23 for following ones.
10. "The national statistical institutes are responsible for selecting the sample, preparing the questionnaires, conducting the direct interviews among households, and forwarding the results to Eurostat (…). The questionnaires are drawn up by each Member State in the national language or languages, taking into account the stipulations made in the Regulation. For every survey characteristic listed in the Regulation, a question or series of questions exists in each questionnaire to permit this information to be supplied. There are thirty-one Labour Force Surveys **conducted by the National Statistical Institutes across Europe** and collected and disseminated by Eurostat. Those national LFS:
    - Use the same concepts and definitions,
    - Follow the International Labour Organisation guidelines,
    - Use common classifications: NACE, ISCO-88(Com), ISCED, NUTS,
    - Record the same set of characteristics in each country.
11. Specific to the 2010 UK sample: results for respondents who are not contacted in waves 2 to 5 or who refuse for circumstantial reasons are carried forward from the previous wave if an interview has been carried out in the previous wave. 34% of the interviews were carried out by proxy. Information on proxy responses can be obtained using the PRXREL variable.
12. The LFS uses dependent interviewing, where answers given at the previous wave are available to interviewers.
13. The UK did in 1997 a study on data quality of proxy responses. Among other things, they found that proxies correctly reported the "age of leaving full-time education" 62% of the times (there was a 6% of data loss). This percentage was 63% for qualifications (I assume this refers to level of education). They estimate gross error by multiplying the proportion of error by the proportion of proxy responses in the same (33%--it seems that level of proxy reporting was the same in 1997), coming to a 11% error for both educational variables overall. "For subjects with first degrees, over 80% matched compared with around 30% for those with low grade GCSEs (or equivalent school leaving qualifications) and vocational qualifications. These lower qualifications are particularly vulnerable to misreporting because they relate to obsolete examinations in many cases and we cannot rely on accurate reporting from the subjects themselves. Other research (Bradley et al, 1996) has shown that respondents often fail to tell other members of the household (or the interviewer without a lot of probing) about low level qualifications which they no longer regard as having any value for them."

Microdata are available upon request. Delivery of the data takes up to/on average 10 weeks. The request must be made on behalf of an institution (university, research) and there is a form that needs to be filled out with quite a lot of information.

Questions for experts (see G1) - drafted by Vasja Vehovar, finalized by Ana Villar

1. What is the recommended minimum cell size in case of cell-weighting (or post-stratification)? 30 units? In what circumstances, if any, would it be acceptable to go down to 20 or 10?

2. Does this same rule hold roughly also for the cells of the control margins in the raking and similar methods, such as regression approaches? And does it hold also for implicit inner cells of raking? By control margins we mean the cells that we are trying to match to population proportions, whether these refer to only one variable (say, proportion of males and females) or to combinations of two or more variables (e.g., proportion for low educated women, high educated women, low educated men and high educated men). By implicit cells we mean the cells resulting from crossing all variables used in the raking.

3. What about additional constraints related to the share (%) of the categories of control variables in the population? Some organizations use the criterion that all categories of the control population margins (e.g. the share of certain region) in raking (and similar methods) should contain at least 5% of the population - regardless of (or in addition to) the cell size criterion. In case of cell-weighting or post-stratification this would transfer to limiting the size of all weighting cells to 5% of the population.

4. What is your experience with truncation? What trimming value would you recommend (2, 3, 4, 5, something else)? We have considered a very conservative value (2) but we typically use a more relaxed value for design weights that account for unequal probability of selection (4).

Appendix F: Comments to weighting procedure

Seppo Laaksonen:

- In my understanding the LFS target population is 15-74 old years (at least in Finland) but that of the ESS = 15+ years old. Ineligibility rates are high in older ages and in many other parts due to partial living in the country or even due to illegal living.
- You thus believe LFS that has also high non-response rates and other problems. True values could be available from the up-to-date register but this is also a problem. Some basic statistics figures are however plausible in most countries.
- In one of your example you did not use any interaction but if you are going to do it by three variables, there will be small sample sizes (respondents) in most cells. There is thus the same problem as in post-stratification even your term is not exactly this.

Vasja Vehovar:

- Target population and age is OK in LFS; there are only negligible differences to ESS,
- We actually obtained control interaction gender x age x education from LFS (Eurostat),
- You are right, this should be not called post-stratification (but historically it is as such labelles in DoW of ESS DACE project); I rather call it population weighting. However, for countries without region control, it is in fact post-stratification and personally I believe this good approach if controls are available (also in LFS they typically use, if they have it, interaction region x gender x age group),
- I agree with your comment regarding cutting the weights. We did it very primitively, but will check at least for lower bound.

The whole issue is basically about administration and data preparation/management and controlling country specifics and handling various patterns of missing values (at ESS and LFS control level), which all need to be repeated for many country countries. We thus found it more flexible to do it via routines in R, rather to use specialised calibration software, which would better handle weights and controls. However, with that, the gains would be negligible, but additional work related to administration etc. would be perhaps substantial.

Kaur Lumiste (National Coordinator for Estonia):

- Raking is a simple method, but statistical offices around the world are using calibration, there are even calibration methods that set boundaries on weights like weights cut at 2 and 4. We think that calibration should be favoured over raking.

Vasja Vehovar:

- We agree that calibration method could perform better. However, we assume that in this case the difference in the final weighting results would be negligible. At the same time, the calibration method would require considerably different data management.

Silke Schneider:

- The classification issues with respect to education have been all solved as much as we could (and well enough for a 3-category education scheme I think), and otherwise were documented, especially if ESS deviates from official classification on purpose. So remaining discrepancies would (for lack of other explanations) be largely explained by non-response bias.
  http://www.europeansocialsurvey.org/docs/round5/survey/ESS5_appendix_a1_e03_0.pdf

Appendix G: Seppo Laaksonen on open issues

I am not sure what is expected from me according to Sabine's request and specifically concerning Hungary data from rounds 2 and 4, but I comment something after reading several documents found.

First, it is hard to say why some Hungarian margins are not enough close to 'true margins'. This is an issue examined by each country where should exist best data for that. Some differences look ordinary when trying to get those 'true values.' This is in general a very big issue since we need these true values for the ESS target population that is not at all easy since there are high ineligibility rates and other problems in statistics. For example, I wonder, how you use LFS in this since its target population is not 15+ years olds people living enough regularly in the country. Some 'guestimates' can be obtained but not 'true values.'

The results by the Ljubljana university are interesting and very useful, But as said I wonder how 'true ' are your margin values? A second question what I do not understand is that you say that you are using the 'interaction' margins for post-stratification. But I do not see any interaction margins like gender*age_group*education. Such data are even impossible to get from most countries. I understand that your method is not post-stratification but 'linear calibration' based on maybe raking-ratio methodology. Post-stratification requires to 'calibrate' the initial weights 'conditional to sampling design.' Raking-ratio and other linear calibrations including logit, sinus hyperbolicus etc are straightforwardly calibrating the certain margins to 'tue ones' without taking into account sampling design. So, please do not use the term 'post-stratification.'

A big point for the future is our new sampling design data file that proposes to collect both 'macro' and 'micro' auxiliary variables. The 'macro' variables are like 'true values' or 'true margins' by possible auxiliary variables. If these could be available, the basic calibrations could be done easily. If, in addition, there are 'micro' variables for the respondents and for the non-respondents, respectively, the other options for re-weightings would be easily applied. My suggestion is response propensity weighting plus the basic calibration. This strategy does not require to take care so much about 'incorrect true values' of margins since the only such margins could be used that are good enough. Even education is difficult to correctly obtain for the ESS target population.

Uncut vs cut weights

As our sampling group knows, it is for me difficult to understand 'cut' weights in such a way that the extremely high weights only have been cut, like above 4 or 2. This is not symmetrically correct, since such 'analysis weights' (the average =1) can also be too small. If the maximums are cut to 4, then the below respective minimum values =1/4 =0.25 should be cut, respectively. The weight variation is not the only motivation for 'cut' or 'trim', we should take care about as unbiased estimates as possible. The cutting is always like violation, not any elegant solution. Before cutting, all possible other solutions should have been attempted. Calibration with respective bounding is not either any correct solution.

Comments on open issues (3.Chapter of this document)

General issue:

As said above, I would concentrate more on using 'micro auxiliary variables' but using such 'macros' only that are plausible (very 'true'). It is easier to add 'micro' variables like you propose, and many others. Their quality needs to be good only for the gross-sample data.

Education:

As said also, education level is not either easy, especially since the classification has been changed over the years. But a certain approximate categorization is not hard for 'micro auxiliary' variables.

Age:
I do not know how you can know ineligibility rates by age. These rates vary a lot from one country to the next.

Region:
The problems as above but small biases do not matter.

Missing values:
I thus recommend propensity score weighting plus calibration with some margin variables. Post-stratification cannot be done correctly without specific assumptions.

Weight trimming:
No weight trimming except no other solution found, but if trimmed then both the minimums and the maximums symmetrically. Such 'cut value like 4' is very subjective. Should be an objective trimming only.