

Can we assess representativeness of cross-national surveys using the education variable?

Verena Ortmanns

GESIS – Leibniz Institute for the Social Sciences
Mannheim, Germany

Silke L. Schneider

GESIS – Leibniz Institute for the Social Sciences
Mannheim, Germany

Achieving a representative sample is an important goal for all surveys. This study asks whether education, a socio-demographic variable covered by virtually every survey of individuals, is a good variable for assessing the realised representativeness of a survey sample, using benchmark data. The main condition for this is that education is measured and coded comparably across data sources. We examine this issue in two steps: Firstly, the distributions of the harmonised education variable in six official and academic cross-national surveys by country-year combination are compared with the respective education distributions of high-quality benchmark data. Doing so, we identify many substantial inconsistencies. Secondly, we try to identify the sources of these inconsistencies, looking at both measurement errors in the education variables and errors of representation. Since in most instances, inconsistent measurement procedures can largely explain the observed inconsistencies, we conclude that the education variable as currently measured in cross-national surveys is, without further processing, unsuitable for assessing sample representativeness, and for constructing weights to adjust for nonresponse bias. The paper closes with recommendations for achieving a more comparable measurement of the education variable.

Keywords: education, comparability, cross-cultural surveys, representativeness, sample quality

1 Introduction

How to achieve good survey data quality is an important issue for the whole survey landscape, including official and academic surveys. In addition to reliable and valid measurements, a key criterion for evaluating survey quality is sample representativeness. Commonly response rates are referred to as an important quality indicator for the representativeness of a sample (see Abraham, Maitland, & Bianchi, 2006; Biemer & Lyberg, 2003; Groves, 2006). However, research showed that low response rates do not necessarily lead to nonresponse bias (Bethlehem, Cobben, & Schouten, 2011; Groves & Peytcheva, 2008), so that this indicator alone is insufficient to assess sample representativeness.

Another simple and commonly used approach to evaluate sample representativeness is to compare the data in question to benchmark data by checking descriptive statistics and distributions for core variables (Groves, 2006; Kamtsiuris et al., 2013; Koch, Halbherr, Stoop, & Kappelhof, 2014; Struminskaya, Kaczmarek, Schaurer, & Bandilla, 2014). These benchmark data are often from official sources such as reg-

ister or census data, and it is assumed that they are the “gold standard” regarding representativeness (Bethlehem & Schouten, 2016; Billiet, Matsuo, Beullens, & Vehovar, 2009; Groves, 2006). Following this approach, we speak of a representative sample if the relative distributions of a core set of stable (e. g. socio-demographic) variables in the survey are equal to the relative distributions in the target population (Bethlehem et al., 2011). This focus is justified when looking at large-scale general population surveys using probability based (rather than e. g. quota) sampling methods and best available sampling frames and designs. In European comparative research, in the absence of suitable register or census data of the target population, the European Union Labour Force Survey (EU-LFS) is commonly used as the benchmark for this purpose.

Mostly socio-demographic variables are used for the comparisons between benchmark data and the surveys in question (e. g. Koch et al., 2014; Peytcheva & Groves, 2009; Struminskaya et al., 2014). The age and gender variables are especially suitable due to their high measurement quality and straightforward comparability. However, age and gender are insufficient criteria on their own for judging a samples’ representativeness. Another commonly-used socio-demographic variable is education, which is also covered in almost all surveys (Hoffmeyer-Zlotnik & Warner, 2014; Smith, 1995). In statistical analyses, education is often used as an inde-

Verena Ortmanns, GESIS – Leibniz Institute for the Social Sciences, P.O. Box 12 21 55, 68072 Mannheim, Germany, e-mail: verena.ortmanns@gesis.org

pendent variable to explain, for example, attitudes, beliefs, and behaviours (Kalmijn, 2003; Kingston, Hubbard, Lapp, Schroeder, & Wilson, 2003). It could be a sensitive marker of representativeness: Several studies show that samples in academic surveys contain an education bias; less educated people are often underrepresented in surveys likely due to selective nonresponse (Abraham et al., 2006; Billiet, Philippens, Fitzgerald, & Stoop, 2007; Couper, 2000; Groves & Couper, 1998). Nonresponse bias, which occurs if the characteristic that influences response propensity is also related to the variables we wish to analyse (Biemer & Lyberg, 2003; Kreuter et al., 2010), is thus particularly likely to occur with respect to education. Being able to use the education variable for constructing weights to adjust for nonresponse would thus be highly desirable.¹ However, the comparison with benchmark data becomes much more challenging with the education variable because it is more complex than the gender and age variables, and therefore contains more possibilities for errors on the measurement side (Billiet et al., 2009; Schneider, 2008b).

From previous research we know that the distributions of the education variable for the same country, year, and age-groups between EU-LFS and other survey data are often not equal, even though supposedly coded in the same way. Kieffer (2010) in her analysis focuses on French data from the EU-LFS and the European Social Survey (ESS) from 2002 to 2008, and identified large discrepancies in the distributions for 2002, 2004 and 2006 but smaller discrepancies for the 2008 data. Schneider (2009) shows inconsistencies between data from the ESS, the EU-LFS, and the European Survey of Income and Living Conditions (EU-SILC) for the years 2002 to 2007 for most European countries. Her analysis uses Duncan's Dissimilarity Index for comparing the distributions of the education variable. Ortmann and Schneider (2015, online first), using the same method, find inconsistent educational distributions across four mostly European public opinion surveys: the ESS, the European Values Study (EVS), the International Social Survey Programme (ISSP), and the Eurobarometer. All authors attribute those inconsistencies to inconsistent measurement procedures rather than non-representativeness.

We extend the study by Schneider (2009) by using data from 2008 to 2012, and the study by Ortmann and Schneider (2015, online first) by adding official surveys - the EU-LFS, EU-SILC, and OECD's Programme for the International Assessment of Adult Competencies (PIAAC). The research question of this paper is: Can we use the education variable for assessing the realised representativeness of the samples of cross-national academic and official surveys? If yes, benchmark data could be used for constructing weights to correct for nonresponse bias (Bethlehem & Schouten, 2016; Kreuter et al., 2010).

In order to answer this question, we firstly present the

methodological background on sample representativeness and the measurement of education in cross-national surveys. Then we introduce the data sources, our analysis strategy and Duncan's Dissimilarity Index as our measure of consistency across surveys. Then the results are presented and interpreted with regards to possible sources of inconsistencies using the Total Survey Error (TSE) framework (Groves et al., 2009; Groves & Lyberg, 2010). The paper ends with conclusions and some practical recommendations for achieving more comparable education variables in cross-national surveys.

2 Methodological background

This section is structured by the two dimensions of TSE which distinguish survey errors resulting from problems of representation and measurement. We first clarify how we use the term sample representativeness, and review different methods for evaluating it. Then we describe the challenges of measuring education in such a way that it can be compared across countries and surveys.

2.1 Sample representativeness

A representative sample is important for surveys in order to achieve data that allow statistical inferences about the whole target population (Biemer & Lyberg, 2003). The terms "representative samples" or "sample representativeness" however have many different interpretations (Kruskal & Mosteller, 1979a, 1979b, 1979c). In this paper, we concentrate on the aspect of achieving equal distributions between the surveyed and the target population in large-scale probability based surveys. If a certain group of the population with specific characteristics is less well covered by the survey sample, it is no longer representative of the population and overrepresents some and underrepresents other groups. Those non-observation errors in principle occur through a combination of coverage, sampling, or nonresponse bias (Bethlehem et al., 2011). There are three main methods for assessing sample representativeness: response rates, R-indicators and benchmark comparisons.

The most commonly used indicator for representativeness is the response rate (Abraham et al., 2006; Biemer & Lyberg, 2003; Groves, 2006). Surveys with very high response rates are commonly regarded as representative, if probability sampling methods are employed and respondent substitution is barred, because they imply a low nonresponse rate. The nonresponse rate indicates the upper limit of the possible nonresponse bias. It "refers to the percentage or proportion of sample cases not included in the eventually realised sample, for whatever reasons (refusals, non-contacts, other reasons)"

¹For a discussion of the merits and effectiveness of weighting and weighting techniques see e. g. Bethlehem (2002), Bethlehem et al. (2011), Gelman and Carlin (2002).

(Heerwegh, Abts, & Loosveldt, 2007, p. 3). However, from research we know that low response rates do not *necessarily* lead to a non-representative or biased sample, if nonresponse is random and no bias occurs (Bethlehem et al., 2011; Groves & Peytcheva, 2008). In addition, response rates also ignore errors due to different sampling frames or sampling methods. Therefore response rates alone are an insufficient indicator for evaluating sample representativeness.

A more recently developed set of indicators assessing representativeness of surveys are model-based representativeness measures, such as the R-indicator, partial R-indicator (Bethlehem et al., 2011; Schouten, Cobben, & Bethlehem, 2009), and other balance and distance indices (Lundquist & Särendal, 2013). These indicators compare the set of respondents to a survey to its gross sample, which includes the respondents as well as the nonrespondents (Bethlehem et al., 2011; Schouten et al., 2009). They therefore predominantly assess nonresponse bias while neglecting potential coverage and sampling biases (Nishimura, Wagner, & Elliott, 2016). These sample-based representativeness indicators require auxiliary data for respondents *and* non-respondents. These auxiliary data are usually taken from the sampling frame, e. g. a population register (Schouten et al., 2009). However, information on the education of survey nonrespondents is not available in most sampling frames, except for some countries' population registers, such as in the Netherlands and the Scandinavian countries. Since we wish to look at a much wider range of countries, for which such auxiliary data is not available, we cannot use this approach for assessing the realised sample representativeness.

The third approach uses benchmark data for evaluating the realised sample representativeness. It compares the distributions of selected variables covered by both the data to be evaluated and the benchmark data. The advantages of this approach are firstly its simplicity from a statistical perspective, and secondly the availability of the required benchmark data. Thirdly, coverage and sampling errors are also reflected in benchmark comparisons. However, using this approach requires that the measurement of the variable(s) to be used is comparable. Another disadvantage of using benchmark data is that these data might not be free from (measurement and representation) errors themselves (Groves, 2006; Koch et al., 2014). Typical variables used for this approach are socio-demographic variables (e. g. Koch et al., 2014; Peytcheva & Groves, 2009; Struminskaya et al., 2014) because these are covered in almost every survey and it is assumed that those are usually measured in an equivalent way. Mostly age and gender are used quite often to evaluate the representativeness of a sample, but also education. However, it is well-known that the measurement of education in cross-national surveys is highly complex, which we turn to next.

2.2 The measurement of education in cross-national surveys

In this paper we thus want to figure out whether the education variable is suitable for evaluating the representativeness of a survey sample. To answer the survey question on educational attainment, respondents typically need to identify their highest formal educational qualification in a list of categories. This list is country-specific, because the national elements of the educational system and the names of the qualifications cannot be input harmonised (Schneider, Joye, & Wolf, 2016). The country-specific answer categories have to be mapped into a standard coding scheme before data collection. This approach is called *ex-ante* output harmonisation (Ehling, 2003). Therefore the survey team has to agree on such a standard coding scheme, and clear guidelines and rules have to be defined for developing the country-specific answer categories and the coding procedure (Ehling, 2003; Eurostat, 2006; Eurostat & OECD, 2014). Most comparative surveys agree on some variant of the International Standard Classification of Education (ISCED).

ISCED was designed by UNESCO in the 1970s and revised in 1997 and 2011 (for details on the most recent update, see Schneider, 2013). It was developed in order to facilitate comparisons of country-specific educational programmes for comparative education statistics. Therefore ISCED defines international levels and types of education (UNESCO-UIS, 2006), and education ministries and national statistical institutes map national educational programmes to these levels and types. Since ISCED 97 is used in the surveys analysed in this article, we limit our presentation to ISCED 97. The main levels of ISCED 97 are:

ISCED 0	Pre-primary education (or not completed primary education)
ISCED 1	Primary education or first stage of basic education
ISCED 2	Lower secondary or second stage of basic education
ISCED 3	Upper secondary education
ISCED 4	Post-secondary non-tertiary education
ISCED 5	First stage of tertiary education
ISCED 6	Second stage of tertiary education.

We focus on these seven main levels and ignore the additional complementary dimensions of ISCED 97, because most of the surveys we look at do not use them (see section 3.1).

3 Data and method

In this section, we introduce the data sources and their education variables in the first part. In the second part, the analysis strategy and the indicator of data consistency are described.

3.1 Data and education coding

For our analysis we select those cross-national survey data that permit the construction of a common education coding scheme based on ISCED, i. e. that claim to use ISCED for education coding. Further criteria are the application of random probability sampling, no substitution of respondents, and coverage of a wide range of European countries. This resulted in the selection of seven diverse cross-national survey datasets on a wide range of topics and with very different cross-national set-up and organisation: the EU-LFS (Eurostat, 2008b, 2009b, 2010a, 2011a, 2012a) and the EU-SILC (Eurostat, 2008c, 2009c, 2010b, 2011b, 2012b) as official data, the OECD's PIAAC (OECD, 2016) and the Eurobarometer (European Commission, 2012, 2014) as policy related studies, and three academic surveys: ESS (ESS, 2008, 2010a, 2012), EVS (EVS, 2011), and ISSP (ISSP Research Group, 2013, 2014). We focus on the years 2008 to 2012.

The EU-LFS provides official quarterly household data for monitoring employment and unemployment in the EU. The data are available from the 1970s onwards and cover all European Union countries. As mentioned above, the EU-LFS is used as benchmark data in this study. We only use the spring (second) quarters of the data in our analyses. On average across years and countries, the response rate for 2008 to 2012 is 78% (also due to compulsory participation in 13 of 31 countries, see Eurostat, 2013). Because of the relatively high response rates, we expect less error of non-observations of lower educated respondents in this data, especially when participation is mandatory, than in the academic surveys. What has to be kept in mind is that the EU-LFS allows proxy-reports. Those, in general, raise the response rates, but may also result in measurement errors. With regard to the education variable, the EU-LFS provides a harmonised variable for all countries consisting of 13 categories, thus distinguishing ISCED main levels as well as some elements of sub-dimensions. We expect the coding of the country-specific qualifications into the official ISCED classification to follow the official ISCED mappings, using the basic principles for implementing ISCED formulated by Eurostat (2006, 2008). The harmonisation process of the country-specific education variables takes place in the statistical institutes of the EU member states rather than centrally at Eurostat. In this study we use the EU-LFS as the benchmark data, because of its wide country coverage, probability sampling methods, relatively high response rates, and large sample sizes, supposedly leading to representative data and precise estimates for any given country.

The EU-SILC was launched in 2003 with the aim of providing cross-sectional and longitudinal official micro-data on income, poverty, social exclusion, as well as living and housing conditions in the EU. The average response rate from 2008 to 2012 is around 80%. The EU-SILC also allows proxy-reports. In the EU-SILC, ISCED main levels 5 and

6 are not distinguished. Coding of country-specific education variables into the ISCED categories for the EU-SILC is also done by the national statistical offices (Eurostat, 2008a, 2009a). Therefore we expect a close match with the EU-LFS data, which was demonstrated for earlier years by Schneider (2009).

OECD's PIAAC data were first collected in 2011/12. This study measures adults' key cognitive skills across OECD countries. The response rate on average is 60%, and there is neither proxy reporting nor compulsory participation in any country (OECD, 2016). PIAAC's education variable adopts the EU-LFS coding scheme and additionally anticipates ISCED 2011 by providing more differentiation at the tertiary level.

The politically-driven Eurobarometer programme was set up by the European Commission in the 1970s to monitor public opinion towards the EU and related topics in all member states. The European Commission unfortunately does not provide information on the response rates of the Eurobarometer studies. Since they do not measure educational attainment on a regular basis, only three Eurobarometer studies from 2010 and 2011 (European Commission, 2012, 2014), which contain main ISCED 97 levels, are included in this study.

The ESS was set up in 2002 and measures individuals' attitudes, beliefs, and behaviour patterns in around 30 European countries. The response rate on average for the years 2008 to 2012 is around 60% (see e. g. ESS, 2014b). Up to 2008, the harmonised education variable consisted of ISCED 97 main levels, but categories 0 and 6 were integrated in categories 1 and 5 respectively. The education variable was changed in 2010, with the aim of achieving more informative and more comparable education variables, introducing a detailed cross-national variable closely anticipating ISCED 2011.

The EVS, which also covers a large number of European countries, was launched in 1981 and also focuses on respondents' values, attitudes, and beliefs. The average response rate for the latest wave (2008) is 56% (EVS & GESIS, 2010). This is the first EVS wave implementing a harmonised education variable representing main ISCED 97 levels.

The ISSP was set up in 1985 and also measures peoples' attitudes and values and extends beyond Europe. For the European countries covered in the ISSP, the response rate on average for 2011 and 2012 is around 50% (see e. g. ISSP, 2015). Before 2011, the ISSP used an education scheme that was specific to the ISSP, but since 2011 one closely related to ISCED has been implemented for measuring educational attainment. Therefore, we will include only ISSP data from 2011 and 2012. However, all upper secondary (ISCED 3) or post-secondary non-tertiary (ISCED 4) qualifications which give access to tertiary education are coded in category 3, and category 4 contains all other upper secondary and post-

secondary non-tertiary qualifications, that are more technically oriented or designed for directly entering the labour market (ISSP – Demographic Methods Group, 2010). Therefore ISSP categories 3 and 4, as well as ISCED levels 3 and 4 of the EU-LFS have to be aggregated to render the coding schemes of both sources comparable.

To summarize, while all these surveys use ISCED 97 as their education coding scheme, each survey defines the specific codes to be used slightly differently. Therefore we further harmonise the different education variables ex-post by focusing on the main ISCED levels, with some adjustments: As the EU-SILC, ESS 2008, and the ISSP do not distinguish between ISCED levels 5 and 6, we combine those two levels. The same is true for ISCED levels 0 and 1, which cannot be differentiated in the ISSP and the ESS 2008 (and many countries in the EU-LFS also fail to make this distinction). The correspondence of the survey-specific ISCED variables and our adapted 5 level version (4 level version for the ISSP) used for the analyses in this study is shown in Table A1 and Table A2 in the appendix.

From each survey, respondents aged 25 to 64 are selected to render samples as comparable as possible. Data are weighted using design weights when available. Cases with missing values on the education variable are excluded from the analysis. This is unproblematic because item-nonresponse on the education variable is generally very low.

3.2 Analysis strategy and method

Our analysis consists of two steps. Firstly, we compare the distributions of the education variable across surveys to see whether the data are consistent across data sources. Secondly, we examine those cases revealing the largest inconsistencies to find out whether these can be explained by differences in measurement procedures, or by lack of representativeness of the sample.

In the first step, for measuring the inconsistencies of the harmonised education variable, we compare the education distributions for the same country and year between the EU-LFS and each other survey presented in section 3.1, by calculating Duncan's Dissimilarity Index (O. D. Duncan & B. Duncan, 1955).² Originally, Duncan's Dissimilarity Index was developed for measuring residential segregation, but it can also be used to measure differences in the distributions of categorical variables more generally. Formally, if x_i denotes the size of category i out of k ISCED categories for country A in year B in survey S, and y_i denotes the same for country A in year B in survey T, the index is defined as:

$$D = \frac{1}{2} \sum_{i=1}^k |x_i - y_i|$$
 We rescale the index to range from 0 to 100 in order to interpret the resulting number as the percentage of cases that would have to change categories in order to achieve an equal education distribution across the two data sources. This can be regarded as the TSE with respect to the

education variable.

In the second step, for cases revealing specifically large deviations, we examine whether those are likely to be caused either by measurement errors in the education variable, or by errors of non-observation, or both. For this analysis we try to unpack the overall discrepancies. To do this, we have a closer look at the frequencies of the ISCED variable across the two surveys in question and check whether the inconsistencies are concentrated in specific ISCED levels or whether they are spread across the education spectrum. If we identify an inconsistency in specific ISCED levels, we have a closer look at the country-specific questions and showcards of the survey (if available) and analyse the exact wording of the categories on the showcard in comparison with the respective information for the EU-LFS. We then check to which ISCED levels the qualifications are coded, and whether this coding appears to differ from the official (EU-LFS) coding. For interpreting the coding in the EU-LFS we used the ISCED mappings of 2013, which contain ISCED 1997 codes used in the EU-LFS, as earlier versions are not publicly available. For the ESS, EVS, ISSP and Eurobarometer, the country-specific education variables and the ISCED variable are included in the datasets, so a simple cross tabulation can be made to identify the mapping used. If we do not find any explanation on the measurement side for the inconsistent education distributions, i. e. if the instrument and coding appear equivalent, we conclude that the representativeness of the sample is probably in question.

One challenge is that it is very difficult to disentangle, let alone quantify, measurement and representation errors empirically. Another challenge when comparing the survey data in question with data from official surveys is that the latter are also not free from errors: The variables of interest could be measured differently, e. g. by allowing proxy respondents, or the samples' characteristics regarding coverage and nonresponse may be different, which both leads to discrepancies in the distributions (Billiet et al., 2009; Groves, 2006). We are aware of the fact that these errors also occur in our benchmark data, the EU-LFS, and that register or census data would be better, if they existed in a comparable fashion across Europe. However, for the reasons mentioned above, this is the most adequate benchmark for this task. Rather than naively assuming the EU-LFS as a "golden standard", when presenting and discussing the results we will try to take potential quality issues with this benchmark data itself into account.

²In the case that some countries run their fieldwork a year later than foreseen (for example Italy and Finland in 2009 instead of 2008 in the EVS), we stick to the main survey year (in this case 2008). We do not expect a substantial change in the distribution of the education variable across two consecutive years because the actual educational distribution in the population only changes very slowly through cohort replacement.

4 Results

In line with the two steps of our analysis strategy, we first present the results regarding Duncan's Dissimilarity Index, with which we identify inconsistencies in the education distributions within countries and years across surveys. We then move on to examine more closely several examples of countries and survey years with large inconsistencies.

4.1 Comparing distributions of the education variable across surveys

For a first overview, Figure 1 shows the boxplots of Duncan's Dissimilarity Index across countries in percent for comparisons between the EU-LFS data and the other six surveys for all possible time points in the years 2008 to 2012. Detailed results for individual countries can be found in Table A3 in the appendix. Comparing EU-LFS and EU-SILC, the median across countries of Duncan's Index is between 4 and 5% in years 2008 to 2012. On average, around 4% of the respondents would have to change into another category to reach equal education distributions across the two datasets. The highest inconsistencies, on average over the five years, are observed for Iceland (16%), Switzerland (15%), and Luxembourg (13%). The smallest deviations between EU-LFS and EU-SILC can be found for the Czech Republic (less than 1%), Slovenia and Austria (both around 2%). The education distributions in these latter countries thus lie very close together which means they are almost perfectly consistent across the two surveys.

When comparing data from PIAAC and EU-LFS, the median of Duncan's Dissimilarity Index is 8%. For Norway (14%), England and Northern Ireland³ (12%), and the Slovak Republic (11%) the highest discrepancies are found. For Austria (2%), France and Sweden (both around 3%) the inconsistencies are smallest.

The median of Duncan's Index in the three Eurobarometer studies in 2010 and 2011 and the EU-LFS of the equivalent years is much higher, between 14 and 18%. We found very high discrepancies between the education distributions of the two data sources of around 40%, averaged over the three comparisons, for the Netherlands, Malta, and Hungary. Small inconsistencies are identified for Slovenia, the Slovak Republic (both around 4%), and Poland (5%).

Comparing ESS 2008, 2010 and 2012 education distributions with those from the corresponding years of the EU-LFS, the median value of Duncan's Index lies between 9 and 11%. High inconsistencies can be found for the UK (25%), Poland (23%), and Denmark (19%) across the three years. The smallest deviations are observed for Bulgaria (3%), Switzerland and Portugal (both around 4%).

With respect to the comparison of EVS and EU-LFS 2008, the median value of Duncan's Index across countries is 14%. We found the largest discrepancies between the two edu-

cation distributions for Estonia (35%), Finland (30%), and Slovenia (27%) and the smallest for the Czech Republic (3%), the Slovak Republic (4%), and Bulgaria (5%).

Finally, comparing the ISSP and the EU-LFS 2011 and 2012, the median value for the inconsistency of the further aggregated ISCED classification (see Table A2 and Section 3.1) also amounts to 14%. On average, the highest discrepancies are observed for Austria (50%), followed by Denmark (33%), and the Slovak Republic (32%). The lowest inconsistencies are found for Latvia (1%), Bulgaria and Portugal (both around 5%).

The overall median inconsistency of education distributions between the six surveys and the EU-LFS for the time period 2008 to 2012 and across countries lies around 10%. The lowest – and substantively non-problematic – median inconsistencies and also the smallest range as well as interquartile range across countries can be observed between the EU-LFS and EU-SILC and the EU-LFS and PIAAC. The large range between EU-LFS and EU-SILC 2011 is the effect of one outlier, namely Iceland. Intermediate median inconsistencies and respectively intermediate ranges and interquartile ranges are identified for the ESS as compared to the EU-LFS. For the comparison of the EU-LFS and the EVS we find a slightly higher median and a higher interquartile range than for the comparison of EU-LFS and ESS data, while the range of inconsistencies is similar. For comparing the EU-LFS with the ISSP (both years) and the Eurobarometer 2011 respectively, we identified the same median inconsistencies. The interquartile range however shows a larger variation of inconsistencies for these benchmark comparisons. For the ISSP 2012 we observe the largest range, caused by the outlier Austria. We find the highest discrepancies when comparing the data of the EU-LFS and the two Eurobarometer studies for 2010, which however show a somewhat lower interquartile range than the comparison with the Eurobarometer 2011 (at constant range).

Overall, the inconsistencies and the interquartile ranges shown in the boxplots vary quite strongly across survey programmes for the same countries and time points. Since the actual education distribution in the population only changes very slowly through cohort replacement, the observed inconsistencies must be ascribed to methodological factors. This may mean either a problem of poor representativeness, or systematic differences in the measurement of education between the surveys. In the next step of the analysis, we will try to disentangle these two factors.

³For this comparison, Scotland and Wales were excluded from the EU-LFS because they are not covered in PIAAC.

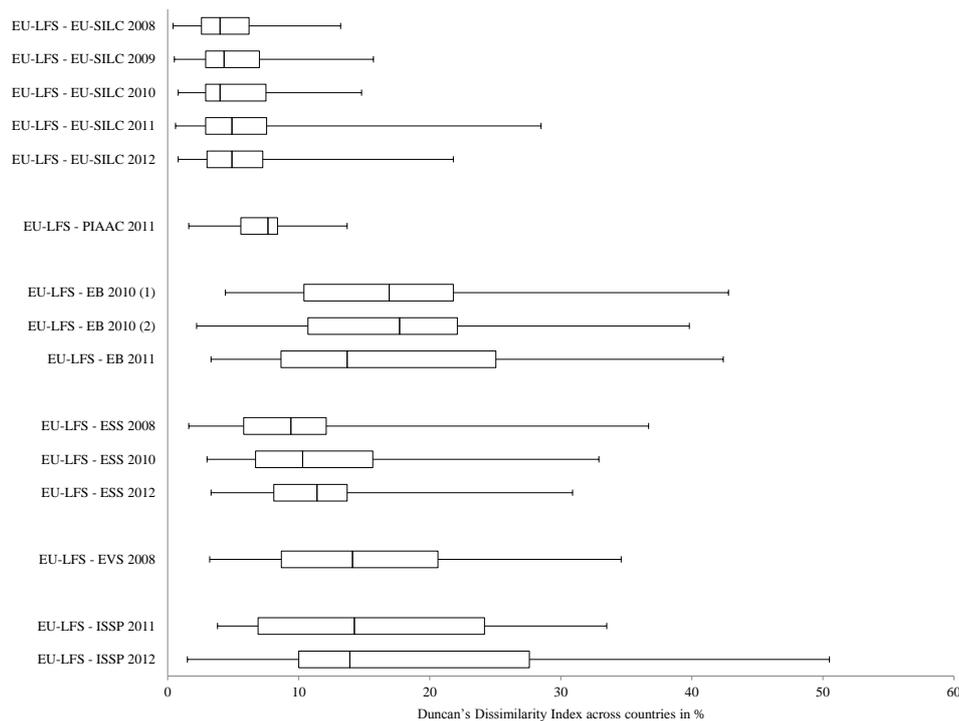


Figure 1. Boxplot of Duncan's Dissimilarity Index across countries for all survey comparisons

Only respondents aged 25-64 for all surveys.

EU-LFS 2008–2012: Eurostat (2008b, 2009b, 2010a, 2011a, 2012a, second quarter, variable HATLEVEL, weighted using variable coeff); *EU-SILC 2008-2012*: Eurostat (2008c, 2009c, 2010b, 2011b, 2012b, variable PE040, weighted using variable PB040); *PIAAC 2011*: OECD (2013, variable edcat7, analysed using complex weight with the International Database Analyzer); *EB 2010, 2011*: European Commission (2012, variable v362), European Commission (2014, variable v105), data weighted to correct regional oversampling for Germany and the UK; *ESS 2008–2012*: ESS (2008, variable edulv1a) ESS (2010b, 2012, variable edulv1b), data weighted using variable dweight; *EVS 2008*: EVS (2011, variable v336, data weighted to correct regional oversampling for Belgium, Germany, and the UK); *ISSP 2011-2012*: ISSP Research Group (2013, 2014, variable DEGREE, data weighted to correct regional oversampling for Germany)

4.2 Explaining inconsistencies between education distributions across surveys

As main factors for explaining the inconsistencies, we distinguish between the two dimensions of the TSE - the measurement and the representation sides, where measurement includes instruments and data processing (Groves et al., 2009). We attempt a deeper interpretation of the results for those 35 country-survey-year-comparisons (affecting 18 countries) showing inconsistencies in the education distributions of more than 25% in at least one comparison between one of the six surveys and the EU-LFS. For each of these comparisons, we first check whether we can find signs of systematic errors on the measurement side, i. e. problems regarding measurement instruments, the response process and

data processing, which in the case of comparative education measurement refers to output harmonisation.⁴ Then, especially if we do not find any hints at measurement and harmonisation problems, we look out for signs of errors of non-observation, especially selective nonresponse. In the follow-

⁴ Some (especially Nordic) countries in some surveys (mostly the EU-LFS) obtain socio-demographic data from population registers rather than by actually asking respondents. This could explain part of the inconsistency found for these countries between the education distribution in the EU-LFS and the other data sources, which are however rather small (apart from Iceland, see below). While register-based data do not contain survey measurement error, we cannot be sure about the quality either; for example they may be incomplete with regards to the education of migrants, and of nationals who have completed their education abroad.

ing section we present one case per survey error source in more detail. Illustrative results for these selected cases can be seen in Figure 2.

Errors related to measurement instruments. The first set of problems that may explain inconsistencies in measured education distributions results from inconsistent or sub-optimal wording of education questions and response categories as well as missing response categories. A rather rare example for *different question wording* (or even choice of different empirical indicators across surveys), which just misses the 25% criterion, is observed for Slovenia in the ISSP, where the education question asks for the last school that was completed rather than the highest educational qualification obtained. While, these two indicators probably correlate highly and this issue thus only explains part of the discrepancy between ISSP and EU-LFS, it is a remarkable lack of input harmonisation.

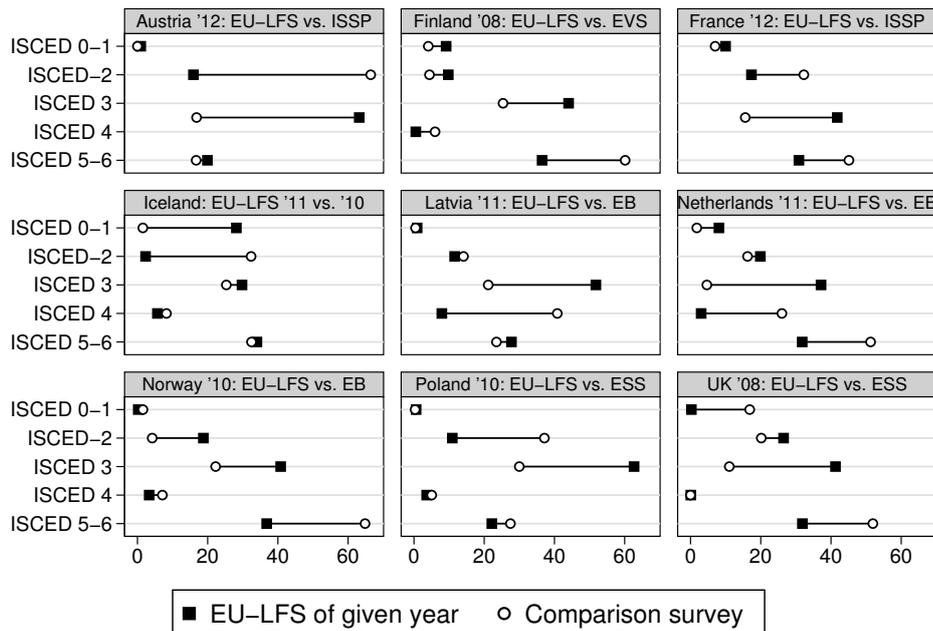
A further example related to the measurement instrument is the *number of questions* asked on the highest educational attainment. Regarding Germany most surveys ask two questions, one on the school leaving certificate, and one on post-school education. Therefore German respondents are used to individual showcards for the school certificates and vocational and higher education qualifications. In the Eurobarometer only one question is asked and consequently only one (but therefore very long) showcard is presented. This could lead to stronger primacy effects in the Eurobarometer, if respondents select the first matching entry, likely a school-leaving certificate, rather than the highest one (as they should). This could likely explain the large discrepancy which is 24% on average between the three EU-LFS and Eurobarometer studies. The largest deviations are observed for ISCED categories 2 and 3.

An example for the use of *vague* or *ambiguous terms* in the questionnaires and on the showcards is France in the ISSP 2012. The inconsistency in the education distributions compared to the EU-LFS 2012 amounts to 29%. Around one third of the respondents are coded into ISCED level 2 in the ISSP whereas in the EU-LFS only 17% fall into this category. Regarding the combined ISCED levels 3 and 4, 16% of the ISSP respondents and 42% of the EU-LFS respondents are coded here (see Figure 2). The ISSP showcard contains ambiguous terms and descriptions rather than specific names of educational qualifications, especially regarding vocational upper secondary and tertiary education. For example, it generically mentions “vocational training after lower secondary school” (“Enseignement professionnel après le collège”) in two response categories. Such terms do not easily correspond to the specific names of French vocational training certificates, programmes, or institutions that respondents may have in mind, such as CAP (“Certificat d’aptitude professionnelle”) or BEP (“Brevet d’études professionnelles”). This could be confusing for respondents

who may not find their specific qualification on the showcard and thus may be unsure which category to pick. The EU-LFS showcard is more precise through offering these specific qualifications as response options. However, the discrepancy at ISCED levels 5 and 6, into which 45% of the ISSP respondents and 31% of the EU-LFS are classified, cannot easily be explained by measurement error because the way the categories are worded should lead to underreporting rather than overreporting of tertiary education in the ISSP. Here we rather think of an education bias in the sample of the ISSP. This probably is in line with the large deviation of nearly 50 percentage points in the response rates: 37% in the ISSP in contrast to 85% in the EU-LFS. Further examples of this kind where we find a mix of showcard issues and selective nonresponse are Denmark, the Netherlands and Sweden in the ISSP. y

A related kind of measurement error occurs when an answer category is entirely missing on the showcard. In this situation, some respondents do also not know which category to choose, but here there basically is none that would really fit. This probably happened in Latvia in the three Eurobarometer studies, where the inconsistency between the Eurobarometer and the EU-LFS data is above 30%. The largest discrepancies are observed for ISCED levels 3 and 4. In the Eurobarometer, more than one third of respondents chose one specific response category that is coded to ISCED level 4, while in the EU-LFS only around 8% fall into ISCED level 4. This category on the Eurobarometer showcard, translated into English, reads “Post-secondary education including professional continuing education, but not higher education programmes (1–3 years after general upper secondary school)”.⁵ Due to the absence of a category referring to vocational training after *lower* secondary school (“pamatskola”) on the Eurobarometer showcard, all respondents having vocational training probably pick this category, whether or not they have actually completed general upper secondary education. In contrast to the Eurobarometer, the showcard of the EU-LFS in Latvia contains five categories covering professional programmes that respondents can more easily choose from and will thus be correctly coded in ISCED. A similar problem regarding missing vocational education categories in the Eurobarometer is observed for upper secondary education in Sweden. Such sub-optimal provision for vocational education is quite common in education questions. This may have several reasons: firstly, vocational education may not be

⁵Name of the Latvian category on the Eurobarometer showcard: “Pēcvidējā izglītība ieskaitot profesionālās tālākizglītības programmas, bet ne augstākās izglītības programmas (1-3 gadi pēc vidusskolas)”. Normally, education answer categories contain country-specific names of educational qualifications, which cannot be translated. Since the Latvian showcard in the Eurobarometer here only provides a generic description, translation is possible in this case.



For Austria and France: Markers between "ISCED 3" and "ISCED 4" refer to "ISCED 3-4"

Figure 2. Education distributions (in percent) in different surveys and years for selected countries

Only respondents aged 25-64 for all surveys. Second quarter data for all data from the EU-LFS data. Figure shows proportions of variable HATLEVEL weighted by variable coeff for data of the EU-LFS.

Austria, France: Eurostat (2012a), ISSP Research Group (2014, variable DEGREE); *Finland:* Eurostat (2008b), EVS (2011, variable v336); *Iceland:* Eurostat (2011a), Eurostat (2010a); *Latvia, Netherlands:* Eurostat (2011a), European Commission (2014, variable v105); *Norway:* Eurostat (2010a), European Commission (2012, variable v362); *Poland:* Eurostat (2010a), ESS (2010b, variable edulvlb, weighted using dweight); *United Kingdom:* Eurostat (2008b), ESS (2008, variable edulvla, weighted using dweight)

considered as formal education; secondly, it may be regarded as irrelevant when the measure of education is only meant as a proxy for academic skills; and thirdly, the number of respondents who have vocational qualifications is estimated to be rather small. All these reasons are problematic in the context of cross-national surveys when different surveys and countries may opt for different solutions in the absence of clear guidance.

Errors related to data processing. Inconsistent application of ISCED, "accidental" or intended, is another important source of inconsistent education distributions on the measurement side. If we find documentation on a decision to deviate from the official ISCED mapping or we find straightforward reasons such as educational reforms, we call this an intended deviation – which should likely not be called an error in the survey in question, but an error or gap in the official ISCED mappings: such deviations are typically made in order to improve comparability across countries and time. This latter situation can only occur in academic surveys because official surveys are bound to use the official ISCED

mappings. We thus define misclassifications as "accidental" when we could not find "obvious" errors or documentation showing why a certain qualification is coded into a different ISCED category than suggested by the official mappings – we then have no reason to think that the deviation was intended.

Firstly, an example where we identify a processing error in the benchmark data of the EU-LFS: Iceland. This inconsistency is identified because in 2011, Iceland in the EU-LFS produces large discrepancies compared to *all* other surveys. Therefore we have a look at the EU-LFS data over time and spot a high value of Duncan's Dissimilarity Index of 33% comparing EU-LFS data of 2010 and 2011. It seems that a large number of respondents previously coded in ISCED level 2 was downgraded to the combined category of ISCED level 0 and 1 in 2011. The coding in 2010 seems to be correct and is also implemented in the other surveys. We could not identify the reason for the shift of coding in the EU-LFS in 2011. Maybe it is "just" a coding error made by a human that slipped through any quality check. This example shows that

our benchmark data is not free from errors, either.

Another factor which may lead to “accidental” misclassification is complications in the communication process between the different teams working on the survey. This may be the explanation for the deviation of around 50% for Austria in the ISSP 2012 from the EU-LFS – this is the highest discrepancy identified in the whole analysis. The largest deviation is on ISCED level 2, in which 16% of the respondents in the EU-LFS and 66% of the respondents in the ISSP are found. For ISCED level 3, the distributions are the other way round. What probably happened is that Austria still used the coding scheme of previous ISSP rounds, in which vocational upper secondary school (“berufsbildende mittlere Schule”), was coded to category 2 instead of category 4 (which is where vocational ISCED 3 qualifications are found in the ISSP since 2011, see section 3.1) as it now should be. Austria did not participate in the ISSP in 2011 and thus may have missed instructions on the changes of the education variable.

A third example of an “accidental” misclassification where we could also identify the specific coding error relates to the Netherlands in the Eurobarometer. The overall discrepancy between the Eurobarometer and the EU-LFS 2011 for the Netherlands is over 40%. In the Eurobarometer around one fourth of the respondents are found in ISCED level 4, whereas only 3% of the EU-LFS respondents belong to this category. Instead, around 37% of the EU-LFS respondents and only 5% of the respondents of the Eurobarometer are coded to ISCED level 3. This can be explained by the misclassification of the school-leaving certificates of upper secondary institutions such as the VWO (“Voorbereidend wetenschappelijk onderwijs”), HBS (“Hogereburgerschool”), and the vocational qualification MBO (“Middelbaar beroepsopleiding”). These qualifications are classified into ISCED level 4 in the Eurobarometer instead of ISCED level 3, as they should be according to the official ISCED mappings. The discrepancy at ISCED levels 5 and 6, into which half of the Eurobarometer respondents and 32% of the EU-LFS are classified, cannot be explained by differences between instruments or processing error. Here we assume an education bias in the Eurobarometer sample. Further examples of “accidental” misclassifications, which are not discussed here in detail, can be found for Hungary in the Eurobarometer and the ISSP (see Ortmann & Schneider, 2015, [online first](#)), Slovenia in the ISSP and EVS, Sweden and the Slovak Republic in the ISSP, and Spain in the Eurobarometer.

Intended deviations from the official ISCED coding are a further possible explanation for some discrepancies, which are however well documented only for the ESS data since round 5 (ESS, 2010a). For Poland we found inconsistent data with Duncan’s Dissimilarity Index of more than 30% between EU-LFS and ESS 2010 and 2012. The largest deviation found at ISCED levels 2 and 3: In the ESS 2010, 37% are coded to ISCED level 2 and 30% to level 3, whereas

in the EU-LFS it is 11% and over 60% respectively. This difference is explained by the decision in the ESS to differentiate between the certificate of basic vocational school before and after an educational reform in 1999. Basic vocational school (“Ukończona szkoła zasadnicza zawodowa”) used to start after 7 years of elementary education *before* the reform, while in the *current* system, it starts after 9 years of general education. Before the reform, individuals thus did not complete ISCED level 2 (which typically lasts 9 to 10 years) before entering basic vocational school, but after the reform, they do. This results in ISCED level 2 for the pre-reform vocational qualification, and ISCED level 3 for the post-reform qualification. In the ESS, respondents who achieved the qualification before 2005, when the reform was fully implemented, are therefore coded to ISCED level 2, and all others to ISCED level 3 (ESS, 2010a, p. 59). In the EU-LFS, all respondents with this qualification are regarded as reaching ISCED level 3, although the majority still went through the old system. Such reforms, increasing the duration of compulsory schooling, are invisible in official education statistics, which may, from a political point of view, be quite desirable. A similar case is observed for Estonia in the EVS 2008 where the EVS decided to downgrade the basic vocational training to the lower secondary level (“kutseõpe põhihariduse baasil”).

The UK in the ESS is another example of an intended deviation in data processing and of an overrepresentation of the highly educated. Overall, the inconsistency for the UK between EU-LFS and ESS data is 37% in 2008 and around 20% in 2010 and 2012. Focusing on the comparison of the 2008 data there is a discrepancy on ISCED levels 0 and 1; in the ESS around 17% are coded to this category, whereas in the EU-LFS it is less than 1%. This is explained by the ESS decision to classify respondents who finished compulsory schooling without school-leaving certificate into ISCED level 1 instead of ISCED level 2 as is done in the EU-LFS. The main discrepancy of the UK is however on ISCED level 3, in which 11% of the ESS respondents but 41% of the EU-LFS are classified. This inconsistency is caused by the decision of the ESS to put the General Certificate of Secondary Education (GCSE) into ISCED 2, although it is officially mapped to ISCED 3C (ESS, 2010a). This latter category describes programmes which do not give access to ISCED level 5, but directly lead to the labour market (or to other programmes at ISCED level 3 or 4). These programmes are thus usually vocational. However, the GCSE is a general school leaving certificate awarded at age 16, which does not specifically prepare for direct labour market entry. In order to improve comparability with other western European countries that offer first school-leaving certificates around age 16 at the end of ISCED level 2, GCSE is classified as ISCED level 2 in the ESS (ESS, 2010a; Schneider, 2008a). A further large difference between the two surveys is found at ISCED

levels 5 and 6, where around 30% of the respondents in the EU-LFS but over 50% of those in the ESS are found. We cannot identify a systematic measurement or processing error here, and therefore strongly suspect selective nonresponse by education (or, less likely, sampling frame issues).

From the examples of showcards using ambiguous terms, incomplete sets of response categories, harmonisation problems, poor communication, as well as “accidental” and intended misclassification, we can conclude that the education variable is not consistently measured across surveys. However, most of the measurement errors are processing errors, which could even be corrected ex-post. Then, assessing sample representativeness using the corrected education variables would be possible. If the measurement instruments however are the main “culprit”, this cannot be repaired ex-post.

Errors of non-observation. In some cases, we could not explain the observed inconsistencies of the education distributions even after close examination of the survey instruments and harmonisation routines. Therefore we now look for further factors influencing sample representativeness. These are, for example, differences in *coverage or sampling frames*, different *sampling designs*, different *survey modes*, as well as *selective nonresponse*.

An example where we think sample representativeness is at risk through the sample design and selective nonresponse is Norway, where we find an inconsistency between the EU-LFS (with mandatory participation in Norway) and the Eurobarometer 73.2 of 2010 of more than 30%. The largest deviation occurs at ISCED levels 5 and 6 to which 37% of the respondents of the EU-LFS and 65% of the Eurobarometer respondents are coded. The EU-LFS uses a random sample from the Norwegian central population register. The Eurobarometer, as in most countries, uses a standard random route procedure by which, in principle, a representative sample can be drawn. However, the success of this approach strongly depends on interviewers implementing it correctly. Here interviewers may have systematically avoided poor neighbourhoods, favoured wealthy ones, or substituted unavailable/refusing lower educated respondents by willing and available higher educated respondents. Another explanation could be that lower educated may have refused to participate in the Eurobarometer more often. We unfortunately cannot separate the errors due to sampling design from those due to selective nonresponse, also because for the Eurobarometer, response rates are not published. The high inconsistencies for a substantial number of countries between the Eurobarometer and the EU-LFS data are particularly alarming when considering representativeness, however we also found many education measurement problems (as described above) in the Eurobarometer.

Another factor which can influence the representativeness of a sample by introducing differential nonresponse is the

survey mode. This might explain the high deviation of the education distribution in Finland in the EVS 2008 compared to the EU-LFS of 29%. We found an overrepresentation of higher educated Finnish people in the EVS: over 60% of the respondents stated that they have tertiary education, whereas in the EU-LFS the proportion is 37%. In the EVS 2008, Finland decided to question respondents using a web panel, while all other countries used face-to-face interviews. This Finnish web panel is based on a random selection from earlier telephone or face-to-face samples of which the recruitment criteria are based on figures from Statistics Finland (EVS & GESIS, 2010). However, it seems that this panel is not a representative sample of the Finnish population. In general, web surveys tend to overrepresent highly educated people (Couper, 2000; Dever, Rafferty, & Valliant, 2008).

These examples show that some of the observed inconsistencies are probably caused by errors of non-observation rather than measurement and processing errors. In these cases, we conclude that random route sampling design (via interviewer effects) and selective nonresponse (e. g. if survey modes differ across surveys) might cause the discrepancies, and indeed representativeness is at risk. For those cases it would be possible to design a weighting factor using the education variable based on the distributions of the EU-LFS to correct for the observed inconsistencies, provided we have in fact excluded all measurement sources of the discrepancies of education distributions.

5 Conclusions and recommendations

The aim of this paper was to examine whether the education variable is appropriate for evaluating the realised representativeness of a survey sample. In the first step of the analysis, we detected small median inconsistencies and low variation in the data of EU-SILC and PIAAC as compared with the EU-LFS. We suspect that these surveys use quite similar measurement instruments and coding procedures, as well as state-of-the-art sampling frames and methods. Intermediate median inconsistencies and medium-sized variation are identified when comparing the ESS with the EU-LFS data. Larger median inconsistencies and variation in the distributions are observed for the comparison of EVS, ISSP and Eurobarometer data with the EU-LFS. These could be due to either measurement or representation errors.

Therefore, in the second step, we diagnosed various kinds of measurement errors by having a closer look at the education distributions, measurement instruments and coding (harmonisation) decisions in individual countries, years and surveys with very high inconsistencies between two education distributions. On the measurement side, we find more processing than instrument-related measurement errors, which hints at a potential to correct errors in the data ex-post. Doing so, assessing sample representativeness would become possible. Obviously, these issues imply a lack of substantive

comparability of the education variable (Billiet et al., 2009; Ortmann & Schneider, 2015, [online first](#)). Only for a few cases with large inconsistencies we conclude that these are mostly caused by errors of non-observation alone, so that here the education variable can be used for assessing sample representativeness.

Therefore, there is strong evidence that educational attainment in many cases is not a good variable for evaluating the representativeness of a survey sample. Consequently, the education variable should not be used when designing weights to adjust for nonresponse bias without the necessary measurement comparability checks. The ESS, for instance, since round 5 adjusts education in only three broad education categories to the EU-LFS (Billiet et al., 2009; ESS, 2014a), and they also reversed intended deviations from official ISCED mappings before doing so. From the results of this study we consider this to be quite a suitable solution (which will however result in somewhat less effective nonresponse-adjustment).

One important limitation of this study is that our benchmark data, the EU-LFS, are not free from errors as demonstrated by the example of Iceland. Especially the use of proxy reporting could lead to measurement error in the reference data because proxies may not know the target person's educational attainment well enough. Another limitation of this study is that errors appearing in every survey are not observed, because the value of Duncan's Dissimilarity Index will be unremarkable in this case. Also, we could not systematically examine all survey error sources because some are not observable with our data, for example social desirability bias (Biemer & Lyberg, 2003; Tourangeau, Rips, & Rasinski, 2000). Social desirability could e. g. express itself by respondents reporting the level of education required for their current job rather than their actual level of education. This could upwardly bias respondents' self-reported attainment (Huddy et al., 1997). However, we do not expect that the prevalence of socially desirable responding would differ across the surveys we examine: they are almost all interviewer-administered and thus prone to similar bias. As another example, older respondents may have difficulties remembering their education level. They may also have more difficulty reporting it, especially if formal qualifications have changed over time and the measurement instrument does not mention the names of outdated qualifications explicitly.⁶ We used the same age range across surveys to minimize the impact of such issues on our results. These response effects cannot be studied in detail using quantitative data, but call for more systematic cognitive pretesting of education questions in all countries.

To conclude, we would like to make some recommendations for improving the consistency of education data across surveys, to improve its substantive comparability and to facilitate the use of this variable for checking sample quality and

constructing weights correcting for nonresponse bias. While some of these recommendations address the survey community as a whole and also international official statistics, others can be implemented by each survey directly.

First of all, surveys need good instruments and showcards which avoid the use of ambiguous terms and unspecific, vague wording, or incomplete sets of response categories. The showcards should instead contain the names of educational qualifications, including formal vocational qualifications, or summary terms that are generally understood by respondents and easily codable to ISCED. Therefore, country teams need the ISCED mappings and guidelines for the development of measurement instruments before developing their questionnaire or should adopt existing instruments from other surveys. Also, more research should be conducted comparatively studying educational systems, qualifications, and careers, including vocational ones, with education measurement in mind.

Secondly, we recommend standardising country-specific education response categories and showcards across surveys in order to elicit more similar kinds of measurement errors in different surveys. No instrument will be without measurement error, but it would be good to produce minimal and consistent errors. Such standardised showcards of course need rigorous testing and regular updates to ensure quality.

Thirdly, we recommend more effective quality assurance and control procedures for background variables and their harmonisation in all surveys. Consistency checks such as those described in this article should be standard for a range of socio-demographic variables, so that especially "accidental" misclassifications can be fixed before data release. Regarding the education variable we strongly recommend ex-post corrections of existing data, and improvements of measurement instruments for future data collections, especially for the Eurobarometer and in the ISSP.

Finally, we would like to question the capability of ISCED to ensure substantive comparability of education data in cross-national surveys. ISCED is, during its development and implementation, vulnerable to political influence, chiefly because education ministries or national statistical institutes determine which national qualifications to map to which ISCED level, and in the latter case, statistical institutes don't always seem to act independently in doing so. At the same time, political education targets that are directly related to ISCED, such as the Europe 2020 goal of reducing the numbers of "early school leavers" (i. e. students leaving education with less than ISCED level 3) to below 10% (Eurostat, 2016), provide an incentive to classify educational programmes at ISCED level 3 even though ISCED level 2 may be substantively more accurate in terms of ISCED classifica-

⁶Educational reforms may actually be one reason for using rather vague terms in education questions, the problematic implications of which we discussed above.

tion criteria.

If the international official statistics community does not achieve stricter quality control of national ISCED mappings, the international survey community may need to find solutions that more reliably produce comparable education data for their own purposes. International academic surveys such as ESS, EVS, ISSP, and the Survey of Health, Ageing and Retirement in Europe (SHARE) could agree on one “alternative” ISCED scheme and adjust the official ISCED mappings to optimise comparability over time and space. Thereby, these surveys would be more comparable with each other. If this alternative variable is coded in detail, it would still be possible to also derive the official ISCED variable in order to check sample representativeness by comparing with official data. For such an academic survey version of ISCED, good documentation is required and the recodes to the official version would have to be published. The ESS since 2010 has tried to go down this route with a number of surveys following suit - SHARE, and probably also the EVS 2017.

Following these recommendations, the statistical consistency and substantive comparability of cross-national education data could be greatly improved. The education variable in academic surveys could then reliably be used for evaluating the realised representativeness of survey samples.

Acknowledgements

We are thankful to Carina Cornesse for her helpful advice and suggestions. We also would like to thank the editor and especially the two anonymous reviewers for their questions and remarks which helped us improving this paper. This research was supported by the Leibniz Association through the Leibniz Competition (formerly SAW Procedure); grant number SAW-2013-GESIS-5.

References

- Abraham, K. G., Maitland, A., & Bianchi, S. M. (2006). Nonresponse in the American Time Use Survey. *Public Opinion Quarterly*, 70 (5), 676–703. doi:10.1093/poq/nfl037.
- Bethlehem, J. D. (2002). Weighting nonresponse adjustments based on auxiliary information. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 275–287). New York: John Wiley & Sons.
- Bethlehem, J. D., Cobben, F., & Schouten, B. (2011). *Handbook of nonresponse in household surveys*. Hoboken, New Jersey: John Wiley & Sons.
- Bethlehem, J. D. & Schouten, B. (2016). Nonresponse error: detection and correction. In C. Wolf, D. Joye, T. W. Smith, & Y.-c. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 558–578). London: SAGE Publications Ltd.
- Biemer, P. P. & Lyberg, L. E. (2003). *Introduction to survey quality*. Hoboken, New Jersey: John Wiley & Sons.
- Billiet, J., Matsuo, H., Beullens, K., & Vehovar, V. (2009). Non-response bias in cross-national surveys: designs for detection and adjustments in the ESS. *ASK: Research & Methods*, 18(1), 3–42.
- Billiet, J., Philippens, M., Fitzgerald, R., & Stoop, I. (2007). Estimation of nonresponse bias in the European Social Survey: using information from reluctant respondents. *Journal of Official Statistics*, 23(2), 135–162.
- Couper, M. P. (2000). Web surveys – a review of issues and approaches. *Public Opinion Quarterly*, 64, 464–494.
- Dever, J. A., Rafferty, A., & Valliant, R. (2008). Internet surveys: can statistical adjustments eliminate coverage bias? *Survey Research Methods*, 2(2), 47–62.
- Duncan, O. D. & Duncan, B. (1955). A methodological analysis of segregation indexes. *American Sociological Review*, 20(2), 210–217.
- Ehling, M. (2003). Harmonising data in official statistics: development, procedures, and data quality. In J. H. P. Hoffmeyer-Zlotnik & C. Wolf (Eds.), *Advances in cross-national comparison. A European working book of demographic and socio-economic variables* (pp. 17–31). New York: Kluwer Academic/Plenum Publishers.
- ESS. (2008). European Social Survey round 4 data, data file edition 4.1. NSD (Norwegian Centre for Research Data) – Data Archive and distributor of ESS data for ESS ERIC.
- ESS. (2010a). ESS round 5 (2010/2011) technical report. Appendix a1 education, ed. 4.0. Retrieved from http://www.europeansocialsurvey.org/docs/round5/survey/ESS5_appendix_a1_e04_0.pdf
- ESS. (2010b). European Social Survey round 5 data, data file edition 3.0. NSD (Norwegian Centre for Research Data) – Data Archive and distributor of ESS data for ESS ERIC.
- ESS. (2012). European Social Survey round 6 data, data file edition 2.0. NSD (Norwegian Centre for Research Data) – Data Archive and distributor of ESS data for ESS ERIC.
- ESS. (2014a). Documentation of ESS post-stratification weights. Retrieved from https://www.europeansocialsurvey.org/docs/methodology/ESS_post_stratification_weights_documentation.pdf
- ESS. (2014b). ESS round 6 (2012/2013) technical report ed. 2.2. Retrieved from http://www.europeansocialsurvey.org/docs/round6/survey/ESS6_data_documentation_report_e02_2.pdf
- European Commission. (2012). Eurobarometer 73.2+73.3 (2-3 2010). GESIS Data Archive, Cologne. ZA5236, data file version 2.0.1, doi:10.4232/1.11473.

- European Commission. (2014). Eurobarometer 75.4 (2011). GESIS Data Archive, Cologne. ZA5564, data file version 2.0.1, doi:10.4232/1.11851.
- Eurostat. (2006). *LFS users guide. Labour Force Survey anonymised data sets*. Luxembourg: European Commission.
- Eurostat. (2008a). *Description of SILC user database variables: cross-sectional and longitudinal*. Luxembourg: European Commission.
- Eurostat. (2008b). European Union Labour Force Survey 2008 (EU-LFS). User database, data file version 2013.
- Eurostat. (2008c). European Union Statistics on Income and Living Conditions (EU-SILC). User database, data file version 2008-3.
- Eurostat. (2009a). *Description of SILC user database variables: cross-sectional and longitudinal*. Luxembourg: European Commission.
- Eurostat. (2009b). European Union Labour Force Survey 2009 (EU-LFS). User database, data file version 2013.
- Eurostat. (2009c). European Union Statistics on Income and Living Conditions (EU-SILC). User database, data file version CROSS-2009-4.
- Eurostat. (2010a). European Union Labour Force Survey 2010 (EU-LFS). User database, data file version 2013.
- Eurostat. (2010b). European Union Statistics on Income and Living Conditions (EU-SILC). User database, data file version CROSS-2010-3.
- Eurostat. (2011a). European Union Labour Force Survey 2011 (EU-LFS). User database, data file version 2013.
- Eurostat. (2011b). European Union Statistics on Income and Living Conditions (EU-SILC). User database, data file version CROSS-2011-1.
- Eurostat. (2012a). European Union Labour Force Survey 2012 (EU-LFS). User database, data file version 2013.
- Eurostat. (2012b). European Union Statistics on Income and Living Conditions (EU-SILC). User database, data file version 2012.
- Eurostat. (2013). *Labour Force Survey in the EU, candidate and EFTA countries – main characteristics of national surveys, 2012*. Luxembourg: European Commission.
- Eurostat. (2016). Europe 2020 indicators – education. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Europe_2020_indicators_-_education
- Eurostat & OECD. (2014). Joint Eurostat-OECD guidelines on the measurement of educational attainment in household surveys, version of September 2014. Retrieved from <http://ec.europa.eu/eurostat/documents/1978984/6037342/Guidelines-on-EA-final.pdf>
- EVS. (2011). European Values Study 2008: Integrated Dataset (EVS 2008). GESIS Data Archive, Cologne. ZA4800, data file version 3.0.0, doi:10.4232/1.11004.
- EVS & GESIS (Eds.). (2010). *EVS 2008 method report: documentation of the full data release 30/11/10*. (GESIS-Technical Reports 2010/17). Bonn: European Values Study and GESIS Data Archive for the Social Sciences.
- Gelman, A. & Carlin, J. B. (2002). Poststratification and weighting adjustments. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey non-response* (pp. 289–302). New York: Wiley & Sons.
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646–675. doi:10.1093/poq/nfl033.
- Groves, R. M. & Couper, M. P. (1998). *Nonresponse in household interview surveys*. New York: John Wiley & Sons.
- Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). New York: John Wiley & Sons.
- Groves, R. M. & Lyberg, L. E. (2010). Total survey error – past, present, and future. *Public Opinion Quarterly*, 74(5), 849–879. doi:10.1093/poq/nfq065.
- Groves, R. M. & Peytcheva, E. (2008). The impact of non-response rates on nonresponse bias. *Public Opinion Quarterly*, 72(2), 167–189.
- Heerwegh, D., Abts, K., & Loosveldt, G. (2007). Minimizing survey refusal and noncontact rates: do our efforts pay off? *Survey Research Methods*, 1(1), 3–10.
- Hoffmeyer-Zlotnik, J. H. P. & Warner, U. (2014). *Harmonising demographic and socio-economic variables for cross-national comparative survey research*. Dordrecht: Springer Science + Business Media.
- Huddy, L., Billig, J., Bracciadleta, J., Hoefler, L., Moynihan, P. J., & Publiani, P. (1997). The effect of interviewer gender on the survey response. *Political Behavior*, 19(3), 197–220.
- ISSP. (2015). Study Monitoring Report 2012, family & changing gender roles IV.
- ISSP – Demographic Methods Group. (2010). ISSP background variables guidelines (press release). Retrieved from http://www.gesis.org/fileadmin/upload/dienstleistung/daten/umfragedaten/issp/members/codinginfo/BV_guidelines_for_issp2011.pdf
- ISSP Research Group. (2013). International Social Survey Programme 2011: Health and Health Care. GESIS Data Archive, Cologne. ZA5800 data file version 3.0.0, doi:10.4232/1.11759.
- ISSP Research Group. (2014). International Social Survey Programme 2012: Family and Changing Gender Roles IV. GESIS Data Archive, Cologne. ZA5900 data file version 2.0.0. Dataset for Hungary: ZA5949 data file version 1.0.0. doi:10.4232/1.12022, doi:10.4232/1.12117.

- Kalmijn, M. (2003). Country differences in sex-role attitudes: cultural and economic explanations. In W. Arts, J. Hagenaars, & L. Halmann (Eds.), *The cultural diversity of European unity. Findings, explanations and reflections from the European Values Study* (pp. 311–337). Leiden: Koninklijke Brill.
- Kamtsiuris, P., Lange, M., Hoffmann, R., Schaffrath Rosario, A., Dahm, S., Kuhnert, R., & Kurth, B. M. (2013). Die erste Welle der Studie zur Gesundheit Erwachsener in Deutschland (DEGS1) – Stichprobendesign, Response, Gewichtung und Repräsentativität. *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 56(6), 620–630.
- Kieffer, A. (2010). Measuring and comparing levels of education: methodological problems in the classification of educational levels in the European Social Surveys and the French Labor Force Surveys. *Bulletin de Méthodologie Sociologique*, 107(1), 49–73. doi:10.1177/0759106310369974.
- Kingston, P. W., Hubbard, R., Lapp, B., Schroeder, P., & Wilson, J. (2003). Why education matters. *Sociology of Education*, 76(1), 53–70.
- Koch, A., Halbherr, V., Stoop, I. A. L., & Kappelhof, J. W. S. (2014). Assessing ESS sample quality by using external and internal criteria. Retrieved from https://www.europeansocialsurvey.org/docs/round5/methods/ESS5_sample_composition_assessment.pdf
- Kreuter, F., Olson, K., Wagner, J., Yan, T., Ezzati-Rice, T. M., Casas-Cordero, C., & Raghunathan, T. W. (2010). Using proxy measures and other correlates of survey outcomes to adjust for non-response: examples from multiple surveys. *Journal of the Royal Statistical Society*, 173(2), 389–407.
- Kruskal, W. & Mosteller, F. (1979a). Representative sampling, I: non-scientific literature. *International Statistical Review*, 47(1), 13–24.
- Kruskal, W. & Mosteller, F. (1979b). Representative sampling, II: scientific literature, excluding statistics. *International Statistical Review*, 47(2), 111–127.
- Kruskal, W. & Mosteller, F. (1979c). Representative sampling, III: the current statistical literature. *International Statistical Review*, 47(3), 245–265.
- Lundquist, P. & Särendal, C.-E. (2013). Aspects of responsive design with applications to the Swedish Living Conditions Survey. *Journal of Official Statistics*, 29(4), 557–582. doi:10.2478/jos-2013-0040.
- Nishimura, R., Wagner, J., & Elliott, M. (2016). Alternative indicators for the risk of non-response bias: a simulation study. *International Statistical Review*, 84(1), 43–62. doi:10.1111/insr.12100.
- OECD. (2013). Programme for the International Assessment of Adult Competencies (PIAAC) 2011. Public Use Files: Data file version 1.0.
- OECD. (2016). Technical report of the survey of adult skills (PIAAC). 2nd edition. Retrieved from http://www.oecd.org/skills/piaac/PIAAC_Technical_Report_2nd_Edition_Full_Report.pdf
- Ortmanns, V. & Schneider, S. (2015, online first). Harmonization still failing? Inconsistency of education variables in cross-national public opinion surveys. *International Journal of Public Opinion Research*. doi:10.1093/ijpor/edv025.
- Peytcheva, E. & Groves, R. M. (2009). Using variation in response rates of demographic subgroups as evidence of nonresponse bias in survey estimates. *Journal of Official Statistics*, 25(2), 193–201.
- Schneider, S. (2008a). The application of the ISCED-97 to the UK's educational qualifications. In S. Schneider (Ed.), *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries* (pp. 281–300). Mannheim: Mannheimer Zentrum für Europäische Sozialforschung (MZES).
- Schneider, S. (2008b). *The International Standard Classification of Education (ISCED-97). An evaluation of content and criterion validity for 15 European countries*. Mannheim: Mannheimer Zentrum für Europäische Sozialforschung (MZES).
- Schneider, S. (2009). *Confusing credentials: the cross-nationally comparable measurement of educational attainment*. Doctoral Thesis. Nuffield College, University of Oxford. Retrieved from <https://ora.ox.ac.uk/443/objects/uuid:15c39d54-f896-425b-aaa8-93ba5bf03529>
- Schneider, S. (2013). The International Standard Classification of Education 2011. In G. E. Birkelund (Ed.), *Class and stratification analysis* (pp. 365–379). Bingley: Emerald.
- Schneider, S., Joye, D., & Wolf, C. (2016). When translation is not enough: background variables in comparative surveys. In C. Wolf, D. Joye, T. W. Smith, & Y.-c. Fu (Eds.), *The SAGE handbook of survey methodology* (pp. 288–307). London: SAGE Publications Ltd.
- Schouten, B., Cobben, F., & Bethlehem, J. D. (2009). Indicators for the representativeness of survey response. *Survey Methodology*, 35(1), 101–113.
- Smith, T. W. (1995). Some aspects of measuring education. *Social Science Research*, 24(3), 215–242.
- Struminskaya, B., Kaczmirek, L., Schaurer, I., & Bandilla, W. (2014). Assessing representativeness of a probability-based online panel in Germany. In M. Callegaro, P. J. Lavrakas, J. A. Krosnik, R. Baker, J. D. Bethlehem, & A. S. Göritz (Eds.), *Online panel research: a data quality perspective* (pp. 61–85). West Sussex: John Wiley & Sons.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.

UNESCO-UIS. (2006). ISCED 1997 International Standard Classification of Education: UNESCO-UIS.

Appendix
Tables

(Appendix tables on next page)

Table A1
 Categories and recodes of the education variables across surveys into 5-level version of ISCED97

EU-LFS	EU-SILC	PIAAC	Euro barometer	ESS until 2008	ESS since 2010	EVS
<i>Preprimary and primary or first stage of basic educ.</i>						
0 No formal educ. or below ISCED 1	0 Preprimary educ. or below ISCED 1	1 Primary or less (ISCED 1 or less)	0 Preprimary educ. or first of basic educ.	1 Less than lower secondary educ. (ISCED 0-1)	0 Not completed ISCED level 1 113 ISCED 1, completed primary educ. 129 Vocational ISCED 2C < 2 years, no access ISCED 3	0 Preprimary educ. or none educ. 1 Primary educ. or first stage of basic educ.
11 ISCED 1	1 Primary educ.					
<i>Lower secondary or second stage of basic education</i>						
21 ISCED 2	2 Lower secondary	2 Lower secondary (ISCED 2, ISCED 3C, short)	2 Lower secondary or second stage of basic educ.	21 Lower secondary educ. completed (ISCED 2)	212 General/pre-vocational ISCED 2A/2B, access ISCED 3 vocational 213 General ISCED 2A, access ISCED 3A general/all 3 221 Vocational ISCED 2C ≥ 2 years, no access ISCED 3 222 Vocational ISCED 2A/2B, access ISCED 3 vocational 223 Vocational ISCED 2, access ISCED 3 general/all 229 Vocational ISCED 3C < 2 years, no access ISCED 5	2 Lower secondary or second stage of basic educ.
22 ISCED 3c (< 2 years)	2 Lower secondary educ.					

Continues on next page

Continued from last page

EU-LFS	EU-SILC	PIAAC	Eurobarometer	ESS until 2008	ESS since 2010	EVS
<i>Upper Secondary education</i>						
30 ISCED 3 (without distinction a, b or c possible, ≥ 2 years)	3 (Upper) Secondary educ.	3 Upper Secondary (ISCED 3A-B, C long)	3 (Upper) Secondary educ.	3 Upper Secondary educ. completed (ISCED 3)	311 General ISCED 3 ≥ 2 years, no access ISCED 5 312 General ISCED 3A/3B, access ISCED 5B/lower tier 5A 313 General ISCED 3A, access upper tier ISCED 5A/all 5 321 Vocational ISCED 3C ≥ 2 years, no access ISCED 5 322 Vocational ISCED 3A/3B, access ISCED 5B/lower tier 5A 323 Vocational ISCED 3A, access upper tier ISCED 5A/all 5	3 (Upper) Secondary educ.
<i>Post-secondary non-tertiary education</i>						
43 ISCED 4 (without distinction a, b or c possible)	4 Post-secondary non-tertiary educ.	4 Post-secondary, non-tertiary (ISCED 4A-B-C)	4 Post-secondary non-tertiary educ.	4 Post-secondary non-tertiary educ. completed (ISCED 4)	412 General ISCED 4A/4B, access ISCED 5B/lower tier 5A 413 General ISCED 4A, access upper tier ISCED 5A/all 5 421 ISCED 4 programmes without access ISCED 5 422 Vocational ISCED 4A/4B, access ISCED 5B/lower tier 5A 423 Vocational ISCED 4A, access upper tier ISCED 5A/all 5	4 Post-secondary non-tertiary educ.

Continues on next page

Continued from last page

EU-LFS	EU-SILC	PIAAC	Eurobarometer	ESS until 2008	ESS since 2010	EVS
<i>First and second stage of tertiary education</i>						
51 ISCED 5b	5 1st & 2nd stage of tertiary educ.	5 Tertiary – professional degree (ISCED 5B)	5 First stage of tertiary educ.	5 Tertiary educ. completed (ISCED 5-6)	510 ISCED 5A short, intermediate/academic/general tertiary below	5 First stage tertiary educ.
52 ISCED 5a		6 Tertiary – bachelor degree (ISCED 5A)	6 Second stage of tertiary educ.		520 ISCED 5B short, advanced vocational qualifications	6 Second stage of tertiary educ.
60 ISCED 6		7 Tertiary – master/research degree (ISCED 5A/6)			610 ISCED 5A medium, bachelor/equivalent from lower tier tertiary	
		8 Tertiary – bachelor/master/research degree (ISCED 5A/6)			620 ISCED 5A medium, bachelor / equivalent from upper/single tier	
					710 ISCED 5A long, master/equivalent from lower tier tertiary	
					720 ISCED 5A long, master/equivalent from upper/single tier tertiary	
					800 ISCED 6, doctoral degree	

EU-LFS 2008–2012: Eurostat (2008b, 2009b, 2010a, 2011a, 2012a, variable HATLEVEL); *EU-SILC 2008–2012*: Eurostat (2008c, 2009c, 2010b, 2011b, 2012b, variable PE040); *PIAAC 2011*: OECD (2013, variable edcat7); *EB 2010, 2011*: European Commission (2012, variable v362), European Commission (2014, variable v105); *ESS 2008–2012*: ESS (2008, variable edulv1a), ESS (2010a, 2012, variable edulv1b); *EVS 2008*: EVS (2011, variable v336)

Table A2
Categories and recodes of the education variables in ISSP and EU-LFS into 4-level version of ISCED 97

EU-LFS		ISSP since 2011	
<i>Preprimary and primary or first stage of basic education</i>			
0	No formal education or below ISCED 1	0	No formal education
11	ISCED 1	1	Primary school
<i>Lower secondary or second stage of basic education</i>			
21	ISCED 2	2	Lower secondary (secondary education completed that does not allow entry to university: end of obligatory school but also short programs (less than 2 years))
22	ISCED 3c (Shorter than 2 years)		
<i>Upper Secondary education and post-secondary non-tertiary education</i>			
32	ISCED 3a, b	3 ^a	Upper secondary (programs that allow entry to university)
41	ISCED 4a, b		
30	ISCED 3 (without distinction a, b or c possible, 2 years and more)	4	Post-secondary, non-tertiary (other upper secondary programs toward the labour market or technical formation)
43	ISCED 4 (without distinction a, b or c possible)		
31	ISCED 3c (2 years and more)		
42	ISCED 4c		
<i>First and second stage of tertiary education</i>			
51	ISCED 5b	5	Lower level tertiary, first stage (also technical schools at a tertiary level)
52	ISCED 5a		
60	ISCED 6	6	Upper level tertiary (Master, Dr.)

EU-LFS 2008–2012: Eurostat (2008b, 2009b, 2010a, 2011a, 2012a, second quarter, variable HATLEVEL);
ISSP 2011–2012: ISSP Research Group (2013, 2014, variable DEGREE)

^a ISCED 3B and 4B are included in ISSP DEGREE variable category 4, not 3, which cannot be differentiated in the ISSP. Therefore ISCED 3 and 4 are summarized.

Table A3
Duncan's Dissimilarity Index for educational attainment distributions across surveys and years per country

	SILC-LFS			PIAAC-LFS			EB-LFS			ESS-LFS			EVS-LFS			ISSP-LFS			Weighted mean	
	2008	2009	2010	2010	2011	2012	2010	2010	2011	2011	2012	2008	2008	2010	2010	2011	2011	2012		
	(1)	(2)	(2)	(1)	(2)	(2)	(1)	(2)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)		(2)
AT	3.3	1.5	1.9	1.4	1.4	1.8	1.6	1.6	17.2	14.4	12.3	-	-	-	-	12.9	-	50.5	-	16.3
BE	6.1	6.5	3.9	4.9	6.3	6.3	8.4	8.4	16.3	18.2	8	6.2	12.3	11.1	8.2	8.2	6.9	-	6.9	8.8
BG	4.6	2.7	1.2	1.5	2.0	2.0	-	-	9.6	8.9	8	1.6	3.6	3.5	4.9	4.9	4.7	5.1	4.7	4.8
CH	-	-	-	16.1	16.2	16.2	-	-	-	-	-	4.9	3	3.5	10.8	10.8	6.9	4	6.9	9.1
CY	2.6	3.7	4.5	3.2	4.2	4.2	7.7	7.7	15.4	20.7	24.9	8.1	5.8	4.6	18.9	18.9	-	-	-	11.3
CZ	0.4	0.8	0.9	0.6	0.8	0.8	5.9	5.9	10.4	11.1	7.1	5.8	12.4	13.6	3.2	3.2	24.4	24.6	24.4	9.1
DE	1.4	2.3	3.6	3.1	2.4	2.4	6.9	6.9	25.6	19.3	26.4	9.7	7.9	10.5	10.5	10.5	5	6.7	5	9.8
DK	3.2	2.2	2.6	2.3	6.1	6.1	10.2	10.2	12.0	16.1	11.9	17.2	19.3	20.3	10.7	10.7	33.5	32.2	33.5	14.9
EE	5.3	5.1	4.7	5.8	4.0	4.0	5.5	5.5	21.8	24.6	14.4	11	10.2	11.4	34.6	34.6	-	-	-	15.2
ES	4.0	4.0	3.2	2.9	2.8	2.8	5.8	5.8	22.8	22.1	25.2	10.2	10.1	12.6	20.4	20.4	9.3	10	9.3	12.3
FI	9.8	8.6	7.7	6.5	5.6	5.6	7.6	7.6	17.0	18.5	16.6	5.5	15.2	14.0	29.1	29.1	9.7	11.6	9.7	14.0
FR	4.0	2.9	8.2	8.1	7.6	7.6	2.7	2.7	9.9	8.3	10.0	9.4	11.5	13.3	14.1	14.1	27.7	29.2	27.7	12.0
GB	11.0	9.1	8.0	9.4	9.6	9.6	12.3	12.3	20.8	18.1	19.2	36.7	21.5	17.8	23.5	23.5	20.8	13.9	20.8	17.9
GR	6.5	8.1	8.6	9	7.5	7.5	-	-	12.6	10.7	10.6	19.3	13.4	-	15.2	15.2	-	-	-	12.7
HR	-	-	-	5.9	4.4	4.4	-	-	-	-	-	10.6	7.0	-	6.5	6.5	9.2	11.5	9.2	7.7
HU	3.2	4.4	3.3	4.4	3.1	3.1	-	-	42.8	38.8	36.7	5.5	5.4	3.3	5.2	5.2	-	30.0	-	16.6
IE	5.9	5.0	6.0	7.4	6.2	6.2	8.3	8.3	17.6	19.9	17.2	18.8	7.0	8.9	12.4	12.4	-	12.7	-	11.6
IS	9.8	10.7	8.4	28.5	21.8	21.8	-	-	19.5	17.7	-	-	-	24.1	17.6	17.6	-	19.7	-	16.0
IT	3.1	3.6	3.3	2.9	3.2	3.2	7.8	7.8	16.9	15.7	17.2	-	12.6	16.1	14.1	14.1	-	11.7	-	11.7
LT	3.2	5.4	3.5	5.1	5.7	5.7	-	-	11.4	11.5	13.7	-	18.5	11.4	21.2	21.2	14.4	10.6	14.4	13.1
LU	7.7	15.7	14.8	13.5	15.0	15.0	-	-	8.9	7.4	4.8	-	-	-	7.1	7.1	-	-	-	9.2

Continues on next page

Continued from last page

	SILC-LFS												Weighted mean					
	2008			2009			2010			2011				2012				
	2008	2009	2010	2010	2011	2012	2011	2012	2012	2011	2010	2009		2008	2010	2012	2011	2012
				(1)	(2)													
LV	2.5	4.3	4.7	31.6	32.5	35.4	7.6	-	-	20.7	-	1.5	13.3					
MT	-	8.6	7.5	41.6	37.1	39.6	-	-	-	-	-	-	22.6					
NL	1.9	3.0	4.0	38.3	39.8	42.4	12.1	17.9	14.1	23.0	25.7	-	18.7					
NO	2.3	0.5	0.8	33.1	26.1	-	13.7	17.0	13.2	19.5	20.1	20.4	16.6					
PL	13.2	13.1	12.3	4.8	3.3	6.2	4.9	32.9	30.9	14.1	3.8	9.4	11.4					
PT	2.2	3.1	2.9	8.3	7.3	13.1	3.4	3.6	5.7	5.1	6.3	-	5.5					
RO	4.6	3.4	2.3	14.1	12.4	9.3	10.6	-	-	10.1	-	-	9.0					
SE	7.2	6.9	7.2	21.1	23.5	25.3	13.6	9.8	10.4	22.0	24.1	27.6	15.4					
SI	1.4	2.1	1.7	5.8	2.2	3.3	6.0	3.8	3.7	26.6	19.8	23.7	11.7					
SK	5.9	7.0	5.5	4.4	4.0	4.5	8.7	10.4	11.1	3.9	31.3	32.0	11.2					
Median	4.0	4.4	4.0	16.9	17.7	13.7	9.4	10.3	11.4	14.1	14.3	13.9	12.0					
Mean	4.9	5.3	5.1	18.3	17.6	17.2	10.4	11.6	11.9	14.9	15.9	18.4	12.6					
Min	0.4	0.5	0.8	4.4	2.2	3.3	1.6	3.0	3.3	3.2	3.8	1.5	4.8					
Max	13.2	15.7	14.8	42.8	39.8	42.4	36.7	32.9	30.9	34.6	33.5	50.5	22.6					

Cases with dissimilarity index above 25% are mentioned in section 4.2.

EU-LFS 2008-2012: Eurostat (2008b, 2009b, 2010a, 2011a, 2012a, second quarter, variable HATLEVEL, weighted using variable coeff); *EU-SILC 2008-2012*: Eurostat (2008c, 2009c, 2010b, 2011b, 2012b, variable PE040, weighted using variable PB040); *PIAAC 2011*: OECD (2013, variable edcat7, analysed using complex weight with the International Database Analyzer); *EB 2010, 2011*: European Commission (2012, variable v362), European Commission (2014, variable v105), data weighted to correct regional oversampling for Germany and the UK; *ESS 2008-2012*: ESS (2008, variable edulv1a) ESS (2010b, 2012, variable edulvlb), data weighted using variable dweight; *EVS 2008*: EVS (2011, variable v336, data weighted to correct regional oversampling for Belgium, Germany, and the UK); *ISSP 2011-2012*: ISSP Research Group (2013, 2014, variable DEGREE, data weighted to correct regional oversampling for Germany). Only respondents aged 25-64 for all surveys.

^a For PIAAC and EU-LFS only Flanders, excluding Wallonia and Brussels; for ISSP and EU-LFS Flanders and Brussels, excluding Wallonia. ^b For PIAAC and EU-LFS only England and Northern Ireland, excluding Scotland and Wales; for ISSP and EU-LFS excluding Northern Ireland. ^c For PIAAC, DE and AT use age group 25 to 65 instead of 25 to 64. ^d For ISSP, adapted ISCED97_4 level is used (see Table A2).