WP6

Analysis of 2009 EU-SILC national survey samples

Vladimir Katchadourian and Emmanuelle Cambois

INED

Final

EUROPEAN HEALTH & LIFE EXPECTANCY
JA:EHLEIS
INFORMATION SYSTEM

The EHLEIS team comprises:

Jean-Marie Robine, INSERM U988 and U710, Montpellier, France, jean-marie.robine@inserm.fr
Herman Van Oyen, Scientific Institute of Public Health, Brussels, Belgium, Herman.VanOyen@wiv-isp.be
Nicolas Berger, Scientific Institute of Public Health, Brussels, Belgium, Nicolas.Berger@wiv-isp.be
ŠárkaDaňková, Institute of Health Information and Statistics of the Czech Republic, Praha, Czech Republic, dankova@uzis.cz.
Bernard Jeune, University of Southern Denmark, Institute of Public Health, Odense, Denmark, BJeune@health.sdu.dk
Henrik Bronnum Hansen, Department of Public Health, University of Copenhagen, Denmark, Henrik.Bronnum-hansen@sund.ku.dk
Ola Ekholm, National Institute of Public Health, Copenhagen, Denmark, oek@niph.dk
Barbara Hjalsted, Danish National Board of Health, Copenhagen, Denmark, bah@sst.dk
Mikkel Baadsgaard, Economic Council of the Labour Movement, Copenhagen, Denmark, mb@ae.dk
Emmanuelle Cambois, INED (Institut National d'Etudes Démographiques), Paris, France, Cambois@ined.fr
France Meslé, INED, (Institut National d'Etudes Démographiques), Paris, France,mesle@ined.fr
Isabelle Mougenot, University of Montpellier II, Montpellier, France, mougenot@lirmm.fr
Gabriele Doblhammer, Rostock Center for Demographic Change, Germany, doblhammer@rostockerzentrum.de
Jürgen Thelen, Robert Koch Institute, Berlin, Germany, ThelenJ@rki.de
Lars Kroll, Robert Koch Institute, Berlin, Germany, L.Kroll@rki.de
Giorgos Ntouros, Hellenic Statistical Authority, Athens, Greece, Geodouro@statistics.gr
Viviana Egidi, University la Sapienza, Rome, Italy, viviana.egidi@uniroma1.it
Wilma J. Nusselder, Erasmus Medical Center, Rotterdam, Netherlands, w.nusselder@erasmusmc.nl
Caspar Looman, Erasmus Medical Center, Rotterdam, Netherlands, c.looman@erasmusmc.nl
Hendriek Boshuizen, National Institute for Public Health and the Environment, Bilthoven, Netherlands,Hendriek.Boshuizen@rivm.nl
Jan-Willem Bruggink, Statistical Office (CBS), Heerlen, Netherlands, jw.bruggink@cbs.nl
Marten Lagergren, National Board of Health and Welfare (SoS/NBHW), Stockholm, Sweden, marten.lagergren@aldrecentrum.se
Carol Jagger, Newcastle University, United-Kingdom, carol.jagger@newcastle.ac.uk
Chris White, Office of National Statistics, Newport, United-Kingdom, Chris.White@ons.gsi.gov.uk
Tony Fouweather, Newcastle University, United-Kingdom, tony.fouweather@newcastle.ac.uk
Stefaan Demarest, Scientific Institute of Public Health, Brussels, Belgium, stefaan.demarest@wiv-isp.be
Denise Walckiers, Scientific Institute of Public Health, Brussels, Belgium, denise.walckiers@wiv-isp.be
KaatjeBollaerts, Scientific Institute of Public Health, Brussels, Belgium, kaatje.bollaerts@wiv-isp.be
Leila Oumeddour, INED (Institut National d'Etudes Démographiques), Paris, France, leila.oumeddour@ined.fr
Isabelle Beluche, INSERM U710/CRLC, Montpellier, France, isabelle.beluche@inserm.fr
Christine Perrier, CRLC, Montpellier, France, christine.perrier@inserm.fr

Contact EHLEIS:
Jean Marie ROBINE, INSERM
Université Montpellier II / U710 - MMDN
Place Eugène Bataillon-CC105
34095 Montpellier Cedex 05
France

Tel: +33 (0) 467 14 33 85
Fax: +33 (0) 467 14 92 95
Email: jean-marie.robine@inserm.fr

Joint action EHLEIS co-funded by DG SANCO (Agreement number 20102301)

Analysis of EU-SILC 2009 national survey samples

Vladimir Katchadourian and Emmanuelle Cambois

INED

Contents

Introduction

As part of the EHLEIS Joint Action, we are exploring the possibility of producing HLY indicators by socio-economic status to assess socially-related health inequalities in the European Union. An essential step is to evaluate the quality of the data that would be available for use after stratifying the samples by age, sex, and socio-economic status. As described in this report, health-related information is collected from all or part of the initial EU-SILC sample, depending on the country and methodology chosen. Some countries, for example, gather all the information through interview surveys. Other countries use administrative sources to obtain some of the information made available in the EU-SILC databases, particularly concerning households (structure, income, and other characteristics); then a supplementary interview survey in the field is used to collect the remaining information, including on health status. Sample structures thus vary from one country to another. This raises the question of whether the data are sufficiently robust for measuring health by social status, considering the size and representativeness of EU-SILC samples.

Representativeness depends on the factors of participation or non-participation in the survey. If they pertain to health and/or social status, they can alter the nature of the sample and the prevalence of activity limitations in the groups surveyed. They can minimize prevalence if non-participants are more "limited" (refusal/inability to participate in survey) or overestimate it if, instead, non-respondents are less limited (economically active young people who are hard to contact). In sum, two sampling effects operate simultaneously: (1) the selection effect, if non-respondents and non-participants differ significantly from respondents in terms of health or education, and (2) the distortion effect, which can alter prevalence in the total population if the structure of the sample differs significantly from that of the population studied, by shifting the mean value towards that of the most represented category.

The selection effect due to non-participation in EU-SILC is not measurable, when the social status and health of non-participants are not known. The selection effect due to non-response in the final sample can be assessed to the extent that some information, notably on social status, may be contained in the "household" database even if the person did not take part in the individual survey. Furthermore, we can measure distortion by comparing the social structures of the final sample with that of the full EU-SILC database and other European data sources such as the Labour Force Survey.

Our report describes the collection methods and characteristics of the national samples, as used in the 2009 EU-SILC survey for all participating countries. We look first at the total national samples for which information is available in the EU-SILC database, and then examine the national samples that were surveyed and responded to questions about health. We give the response rates and the representativeness of responses relative to the age and social-status structures of the national populations drawn for other sources.

I. The EU-SILC survey: description, collection method, and response rate

The EU-SILC (European Union – Statistics on Income and Living Conditions) programme aims to collect comparable cross-sectional and longitudinal data across the EU on income, poverty, social exclusion, and living conditions. EU-SILC was launched in six Member States in 2003. In its 2009 version, the programme covered 576,208 people across 29 European countries: the EU27 plus Norway and Iceland. The programme is based on the European Statistical System (ESS). The European Commission (Eurostat) provides the overall framework, and the national statistical institutes (NSIs) are in charge of implementation in each country. EU-SILC includes some questions on health, in particular the Global Activity Limitation Indicator (GALI), which serves to calculate the European life expectancy measure called Healthy Life Years (HLY).

1. EU-SILC: collection methods and participation rates

a) From household data files to individual data files

The EU-SILC database comprises four information files: (1) the **household register file**, which provides data on all selected households; (2) the **individual register file**, containing information on all household members; (3) the **household data file**, supplying data on households in which at least one eligible person has responded to both parts of the interview (household part and individual part); (4) **the individual data file**, which provides full data for eligible persons (aged 16+) collected from interviews or reliable external sources. EU-SILC database comprises the two individual data files. The individual participation rates—combining the contact rate, the household response rate, and the individual response rate in each household—vary considerably from one country to another, from the high rates obtained in Cyprus, Romania, and Slovakia (approximately 90% and over), down to the rates observed in Belgium, Norway, Luxembourg, and Denmark, which fall below 70%, and even below 60% in the last two countries.

Table 1: Participation rates, number of persons covered, and percentages of individuals recorded using each collection method, by country[1]

		Individual participation rate	EU-SILC "individual information" sample	Collection method to obtain information available in EU-SILC databases				
				Registers	PAPI	CAPI	CATI	Self-administered
AT	Austria	71.1%	13,610		-	58%	42%	-
BE	Belgium	62.7%	14,721		-	100%	-	-
BG	Bulgaria	77.2%	15,047		100%		-	-
CY	Cyprus	89.5%	9,283			100%	-	-
CZ	Czech Rep.	82.3%	23,302		79%	21%	-	-
DE	Germany	76.5%	28,368		-	-	-	100%
DK	Denmark	53.5%	15,025	x	-		93%	7%
EE	Estonia	74.0%	13,542		2%	98%		-
ES	Spain	81.0%	36,865		-	93%	7%	-
FI	Finland	79.2%	25,157	x	-	4%	97%	-
FR	France	82.7%	25,611		-	100%	-	-
GR	Greece	84.0%	18,035		87%	10%	4%	-
HU	Hungary	84.5%	25,053		100%		-	-
IE	Ireland	78.9%	12,641	x	-	100%	-	-
IS	Iceland	73.1%	8,545	x	-	-	100%	-
IT	Italy	83.7%	51,196		100%	-	-	-
LT	Lithuania	86.9%	12,852		69%	-	31%	-
LU	Luxembourg	51.9%	11,406		100%	-	-	-
LV	Latvia	78.3%	14,403	x	7%	57%	36%	-
MT	Malta	79.8%	10,213		-	100%	-	-
NL	Netherlands	83.4%	23,687	x	-	-	100%	-
NO	Norway	60.4%	13,855	x	-	1%	99%	-
PL	Poland	76.3%	38,541		100%	-	-	-
PT	Portugal	86.4%	13,013		4%	96%	-	-
RO	Romania	96.2%	18,703		100%	-	-	-
SE	Sweden	73.0%	18,441	x	-	-	100%	-
SI	Slovenia	77.7%	29,576	x	-	47%	53%	-
SK	Slovakia	88.5%	16,137		100%	-	-	-
UK	United Kingdom	71.3%	19,380		-	100%	-	-

[1] Eurostat, *2009 comparative EU intermediate quality report*, version 3, July 2012.

Unlike many international surveys, the EU-SILC survey does not rely on a harmonized method for collecting information. We identify four methods in Table 1. Participating countries were allowed to gather the information for the four files through interview surveys of a population sample or by using national statistical sources that meet Eurostat data-quality criteria. Nineteen countries collect nearly all their information from face-to-face CAPIs (Computed Assisted Personal Interviews) or PAPIs (Paper And Pencil Interviews). Three others rely on a combination of face-to-face and telephone interviews. Nine countries (Netherlands, Sweden, Norway, Denmark, Finland, Slovenia, Iceland, Ireland, and Latvia) use data from their national registers for some of the information on households and individuals, supplemented by interviews to gather individual data. Ireland and Latvia collect most of their data from face-to-face interviews. Slovenia uses a 50-50 combination of face-to-face and telephone interviews. The other six countries mainly use telephone interviews. Self-administered questionnaires are used for 7% of interviews in Denmark and for all interviews in Germany.

Individual response rates concern a selected portion of the initial sample for which collective information (i.e., on the respondents' households) is gathered. We observed that the country with participation rates below 70% are generally country for which part of the household/individuals information is collected in population register and the other part by CATI interview. This telephone mode explains part of the non participation. Other methodological issues could be further considered, such as the overall structure of the data collection: in the Netherlands, the EU-SILC individual questions come after the labour force survey questions, which increases the risk of loosing people, due to the length of the interview.

b) From individual data files to the sample including health questions

In the database which supplies individual information, we observe a further attrition when working on a number of individual variables, including health data (Table 2). This new attrition is due to the national survey design for determining the persons eligible to answer these specific questions. First, most individual variables are only collected for the 16 year old and over, whereas the data file contains other demographic information for all household members (including below 16 year old). A large proportion of the missing individual data is actually due to the fact that the under-16s are ineligible for this part of the database. When excluded the age issue, attrition between the EU-SILC individual files and the EU-SILC samples with health questions for four possible reasons reported in Table 2:

- The persons did not respond to the individual questionnaire compiling all the variables studied here and the reason why is not mentioned (unknown issue)
- The person was not contacted available to participate and information was provided by a proxy, preventing the collection of health variables in certain countries. This applies to countries using CATI (Denmark, Finland, Iceland, Netherlands, Norway, and Sweden) and the Czech Republic (Proxy issue),
- Part of the individual sample did not answer a questionnaire, as all the data are taken from national registers. This is the case in Slovenia and, marginally, in Norway (Register data issue)
- The persons answered the individual questionnaire, but not the "Activity limitation" question. This is the least frequent cause, and concerns Hungary, Italy, Lithuania, and Poland more specifically (Non-response issue)

We see in Table 2 that the age issue explains a significant part of this attrition. The proxy issue comes second. Not all countries process data collected by proxy in the same way. Some countries, including most of those collecting information by means of CATI, do not allow proxy response for answering the "Activity limitation" question. This generates a large percentage of missing data, whereas the telephone contact in the household can often be another household member than the one selected at random for the survey. Most of the other countries have chosen to document the variable from interviews with someone other than the one selected. Unfortunately, it is hard to identify the situations clearly, as two variables that allow the identification of proxy interviews seem to be

contradictory in certain cases. In the countries using registers and CATI, people who do not document the “Activity limitation” variable for "proxy response" reasons are not identified as responding by proxy in the “type of interview” variable. In the Czech Republic, people responding "by proxy" for the “type of interview” variable are shown as responding to the GALI question with “missing data” instead of “by proxy" as in the previous case. Slovenia seems to construe the term “proxy” for both variables as the exclusive use of registers for collecting individual EU-SILC data. In the other countries, coding inconsistencies are marginal. In an article analyzing the cross-sectional data of the 2007 survey[2], Maria Lacovou, Olena Kaminska, and Horacio Levy highlight the problem posed by using substitutes for individuals or households who have not been contacted or have refused to respond—a practice that makes the sample less representative. Absent information allowing the identification of substitute persons, the authors recommend excluding the countries concerned from analyses based on national figures. In the 2009 survey, the problem affects Ireland and Spain.

Table 3: EU-SILC and final sample population size, attrition, and causes of attrition by country

	Initial population size and attrition			Distribution of persons lost by cause of loss				
	EU-SILC individual files	EU-SILC samples with health questions	% Attri-tion	Unknown	PROXY	Register data	Non-response	Age under 16
AT	13610	11054	19%	0.0%	0.0%	0.0%	0.0%	18.7%
BE	14721	11651	21%	0.0%	0.0%	0.0%	0.6%	20.2%
BG	15047	13148	13%	0.4%	0.0%	0.0%	0.0%	12.2%
CY	9283	7553	19%	0.0%	0.0%	0.0%	0.0%	18.6%
CZ	23302	16827	28%	0.0%	12.2%	0.0%	0.0%	15.5%
DE	28368	23686	17%	0.5%	0.0%	0.0%	0.4%	15.6%
DK	15025	5866	61%	0.0%	38.6%	0.0%	0.0%	22.4%
EE	13542	11220	17%	0.0%	0.0%	0.0%	0.6%	16.5%
ES	36865	30418	17%	0.0%	0.0%	0.0%	1.1%	16.4%
FI	25157	9962	60%	0.6%	38.5%	0.0%	0.1%	21.3%
FR	25611	20113	21%	0.0%	0.0%	0.0%	0.4%	21.0%
GR	18035	15045	17%	0.7%	0.0%	0.0%	0.0%	15.8%
HU	25053	20354	19%	0.3%	0.0%	0.0%	2.1%	16.3%
IE	12641	9900	22%	0.0%	0.0%	0.0%	0.0%	21.7%
IS	8545	2895	66%	0.0%	40.7%	0.4%	0.1%	25.0%
IT	51196	42159	18%	0.4%	0.0%	0.0%	1.5%	15.8%
LT	12852	10700	17%	3.0%	0.0%	0.0%	0.9%	12.8%
LU	11406	8491	26%	0.3%	0.0%	0.0%	0.8%	24.4%
LV	14403	12066	16%	0.0%	0.0%	0.0%	1.0%	15.3%
MT	10213	8478	17%	1.6%	0.0%	0.0%	0.1%	15.3%
NL	23687	9717	59%	0.0%	35.6%	0.0%	0.0%	23.3%
NO	13855	5349	61%	0.0%	35.2%	1.2%	0.6%	24.4%
PL	38541	29228	24%	0.0%	0.0%	0.0%	6.3%	17.9%
PT	13013	11091	15%	0.5%	0.0%	0.0%	0.0%	14.2%
RO	18703	16282	13%	0.3%	0.0%	0.0%	0.0%	12.7%
SE	18441	7540	59%	0.0%	39.3%	0.0%	0.0%	19.8%
SI	29576	9276	69%	0.0%	0.0%	53.6%	0.0%	15.0%
SK	16137	13636	15%	0.1%	0.0%	0.0%	1.1%	14.3%
UK	19380	15359	21%	0.0%	0.0%	0.0%	1.3%	19.4%

[2] Maria Iacovou, Olena Kaminska, and Horacio Levy, “Using EU-SILC data for cross-national analysis: strengths, problems and recommendations”, Institute for Social and Economic Research, Essex University, 2013.

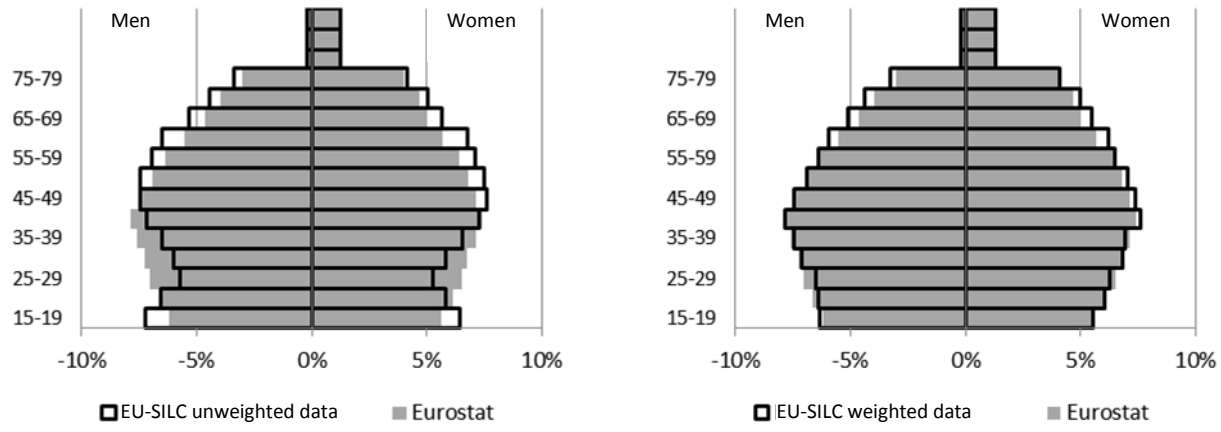
For the seven countries combining registers and telephone interviews, the attrition between the sample of individuals and the sample responding to health questions is massive. Even if strong, this selection effect does not necessarily generate a major bias. Here as well, the representativeness of the final sample will depend on the type of selection responsible for the attrition, analyzed in the following section of the report. As we focus exclusively on the adult population, the missing data for younger persons is arguably not an obstacle to the use of these data. Furthermore, countries provided a weights adjusting for the missing part of the sample. However, these weights might not fully correct for the educational distribution as it is not systematically used as an adjustment criteria and moreover, the weight do not adjust for missing data on specific questions (such as health) when individuals did participate to the other part of the interview (for instance in case of proxy when allowed): this might be linked to health and education. We analyze the nature of the attrition in detail in a later section.

2. Sample size and structure

The national samples of the individual databases range from 8,545 people for Iceland to 51,196 for Italy. Sparsely populated countries generally supply samples that are large relative to their total population. Examples include Cyprus, Iceland, Luxembourg, Malta, and Slovenia, all of which collect information on more than 1% of their national population. By contrast, the largest countries in Europe supply information on the smallest proportions of their population. The prime examples are Germany, France, and the United Kingdom, but their modest relative contributions are offset by the large size of their samples in absolute terms. After performing the stratification according to different socio-demographic variables of interest, we shall examine whether the sample sizes pose a problem for analytical purposes or if the number of people covered suffices, even in the countries with small populations. To assess sample representativeness, we compare their structures by age, education level, and occupational category with the reference figures supplied by another data source from Eurostat namely the labour force survey or the International Labour Organisation. We concentrate on the age groups for which this information is published.

Comparing the EU-SILC individual database with Eurostat data, before and after adjustments using the weights supplied by the countries, we typically observe a slight under-representation of the 25-45s, offset by an over-representation of the over-45s and the 16-19s.

Figure 1: Age distributions of all countries combined, calculated by Eurostat and EU-SILC, with and without weighting



These variations may be due to the greater activity and mobility of the age group that includes students and young workplace entrants—and is thus harder to contact—and to the lesser mobility of older age groups. Weighting methods have not been harmonized at EU level. Age and sex variables have been used in all weight-calibration processes. Some countries also use other variables such as

the size of the municipality of residence or income. Among the variables of interest in this study, education level has been used as an adjustment variable only in Denmark, France, and Slovenia.

We use two cross-sectional weighting variables. The first is compiled for all persons included in the individual database. The second is compiled for the seven countries that combine registers and telephone interviews (DK, FI, IS, NL, NO, SE, and SI). The second variable, compiled only for persons who have been selected and have responded to the survey, serves to adjust the biases due to poor collection of information on certain variables including the health variables.

We shall study in greater detail the similarities or distortions in the sample studied (sample of persons answering the health questions) relative to the population distributions by age, education level, and socio-occupational category.

II. Description of variables of interest and missing data in final sample

Our goal is to study health levels in European countries and the related social inequalities. We can measure social categories from education levels or socio-occupational categories. We shall now describe the variables of interest concerning health ("Self-perceived health", "Chronic health problems", and "Activity limitation"), education ("Highest education level attained"), and occupation ("Occupation"), as well as the ways in which they have been collected. This leads us to select a part of the population recorded in the EU-SILC database, namely, the respondents who have answered questions on the "Activity limitation" variable. They form the final sample.

1. The variables of interest in EU-SILC

Description of health variables

- **Activity limitation**: The Global Activity Limitation Indicator (GALI) encompasses limitations due to illness and physical or mental health problems that have been hampering the respondent's usual activities for at least six months at the time of the survey (the time frame concerns activity limitation, not the health problems that may have preceded it). Three choices are offered: *yes, strongly limited – yes, limited but not strongly – no, not limited*. Eurostat recommends avoiding proxies for this variable, which, as the variable used to calculate HLY, is the focus of our study.

- **Self-perceived health:** This variable consists in the subjective assessment of the respondent's general health, not his or her health status restricted to the time of the interview. Perceived health encompasses the multiple dimensions that can define a person's health status, i.e., his or her physical and emotional condition. The variable is coded with five values: *very good – good – fair – bad – very bad*. Proxies are not allowed for this variable.

- **Chronic health problems:** The respondent is asked about chronic or long-term health problems. A medical diagnosis is not required. The assessment is based solely on the respondents' statements. The variable is coded with two values: *yes - no*. Eurostat recommends avoiding proxies for this variable as well.

"Education" variable for recording highest education level attained

The "Education" variable indicates the highest education level attained. The variable coding follows the seven least detailed aggregation levels of ISCED (International Standard Classification of Education) developed by UNESCO in its 1997 revision and not in its latest revision, introduced in 2011. Only six values have been included in EU-SILC survey data since the survey was launched. The values range from 0 to 5, with 5 combining ISCED levels 5 and 6. For a definition of each level, see the Annex.

Aggregations recommended by Schneider and Müller

Pre-primary education	0
Primary education	1
Lower secondary education	2
Secondary education	3
Post-secondary non-tertiary education	4
First stage of tertiary education	5
Second stage of tertiary education	6

In an article discussing the reliability of the analysis of education levels in Europe based on EU-SILC survey data[3], Silke Schneider and Walter Müller recommend a further aggregation of the six levels into three categories for producing meaningful results, as this would restrict the very wide variations in the categories' population size. Admittedly, the further aggregations limit the analytical usefulness of the "Education" variable, but this has been largely reduced ahead of the analysis. The coding registers information exclusively at the most aggregated levels and makes it impossible to consolidate minor groups on other criteria, such as consolidations for general education and vocational education. Schneider and Müller argue that their three aggregations yield fairly stable categories for the successive survey waves and for making comparisons with the Eurostat EU-LFS (Labour Force Survey). We maintain these aggregations for analyzing the variables.

"Occupation" variable for recording the socio-occupational category

The "Occupation" question codes respondents' occupations according to ISCO-88 (International Standard Classification of Occupations, 1988)[4]. For persons not in employment at the time of the interview, the most recent occupation is recorded. Introduced in 1988 and administered by the International Labour Office, ISCO-88 is one of the standard international social and economic classifications. It is divided into four hierarchical groups. Each has a tree structure and is numbered between 0 and 9. Thus, the first digit designates a *major group*, the second a *sub-major group*, the third a *minor group,* and the fourth a *unit group*. EU-SILC compiles only the first two hierarchical levels. In this study, we look only at the major groups, an aggregation level that provides enough observations for a joint study of occupational status and health status. The codes for the sub-major groups are listed in the Annex.

2. Missing data and final sample

Table 2 shows the percentages of missing data for the "Education", "Occupation", and "GALI" individual variables. The "Education" and "Occupation" variables can be collected at different levels of the questionnaires and can be compiled from the EU-SILC database, including for persons who did not respond to the interview part of the survey. By contrast, the health questions are answered only in the interview part. In the countries combining registers and CATI interviews, if the household member sampled at random for inclusion in the EU-SILC is not the person who responded in the interview—i.e., is a proxy—then the health questions are not asked.

[3] S. Schneider and W. Müller, "Measurement of Education in EU-SILC – Preliminary Evaluation of Measurement Quality", Equasolc Working Paper 2009/5. http://www.equalsoc.org/publication/283, retrieved 21 November 2011.
[4] ISCO-88 (International Standard Classification of Occupations)
http://www.ilo.org/public/english/bureau/stat/isco/isco88/major.htm, retrieved 22/10/2012.

Table 2: Missing data for "Education", "Occupation", and "Activity limitation" (GALI) variables in EU-SILC database by country and age group

	Education			Occupation			Activity limitation		
	30-69	70-79	80 +	30-69	70-79	80 +	30-69	70-79	80 +
AT	0%	0%	0%	4%	1%	2%	0%	0%	0%
BE	1%	0%	1%	1%	0%	1%	0%	0%	1%
BG	0%	0%	0%	1%	1%	1%	0%	0%	0%
CY	0%	0%	0%	0%	0%	0%	0%	0%	0%
CZ	0%	0%	0%	1%	0%	0%	12%	6%	8%
DE	0%	0%	0%	1%	0%	0%	1%	1%	1%
DK	1%	1%	17%	5%	7%	7%	50%	41%	29%
EE	1%	0%	0%	0%	0%	1%	1%	0%	0%
ES	1%	1%	1%	1%	0%	1%	1%	0%	1%
FI	1%	1%	2%	0%	0%	1%	50%	42%	33%
FR	0%	0%	0%	2%	3%	7%	0%	0%	1%
GR	0%	0%	0%	0%	0%	0%	0%	0%	0%
HU	0%	0%	0%	0%	0%	0%	3%	2%	3%
IE	2%	1%	0%	0%	0%	0%	0%	0%	0%
IS	1%	1%	0%	1%	1%	1%	54%	43%	32%
IT	0%	0%	0%	0%	0%	0%	2%	2%	2%
LT	0%	0%	0%	0%	0%	0%	4%	2%	4%
LU	0%	0%	0%	0%	0%	0%	2%	1%	0%
LV	1%	0%	0%	0%	0%	0%	1%	0%	0%
MT	0%	0%	0%	0%	1%	3%	0%	0%	0%
NL	1%	1%	1%	4%	2%	2%	46%	35%	26%
NO	3%	1%	1%	10%	40%	32%	49%	42%	34%
PL	6%	4%	6%	0%	0%	0%	6%	4%	6%
PT	0%	0%	0%	0%	0%	0%	0%	0%	0%
RO	0%	0%	0%	0%	0%	0%	0%	0%	0%
SE	0%	1%	4%	3%	3%	4%	50%	41%	27%
SI	0%	0%	0%	4%	1%	2%	61%	53%	55%
SK	0%	1%	0%	0%	0%	0%	2%	1%	1%
UK	9%	3%	5%	2%	2%	1%	1%	1%	0%
Total	1%	1%	1%	1%	1%	1%	14%	8%	7%

Denmark has high proportions of missing data for the "Education" variable for persons over age 80, as do Poland, Sweden, and the United Kingdom, to a lesser extent. Some countries exhibit a slightly above-average proportion of missing data before age 80 (NO, PL, UK). The United Kingdom has 9% missing data in the 30-69 age group. The "Occupation" variable is less well documented than the "Education" variable. The very high proportion of missing data for the GALI variable is largely due to interview-related attrition, particularly in countries that conduct interviews by telephone (Denmark, Finland, Iceland, Netherlands, Norway, Sweden, and Slovenia). The missing data decrease with age, reflecting the fact that telephone surveys are more likely to miss persons in the youngest age groups.

By concentrating on the extraction of data from the EU-SILC database to document the GALI variable—our variable of interest—we obtain a better response rate for the "Education" and "Occupation" variables (Table 4).

Table 4: Missing data for "Education", "Occupation", and health variables in final sample ("Activity limitation" documented), by country and age group

	Education			Occupation			Self-perceived health			Chronic health problems		
	30-69	70-79	80 +	30-69	70-79	80 +	30-69	70-79	80 +	30-69	70-79	80 +
AT	0%	0%	0%	4%	1%	2%	0%	0%	0%	0%	0%	0%
BE	0%	0%	0%	1%	0%	1%	0%	0%	0%	0%	0%	0%
BG	0%	0%	0%	1%	1%	1%	0%	0%	0%	0%	0%	0%
CY	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
CZ	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%
DE	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%
DK	1%	1%	**18%**	2%	1%	2%	0%	0%	0%	0%	0%	0%
EE	0%	0%	0%	0%	0%	1%	**20%**	**12%**	**19%**	0%	0%	0%
ES	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%
FI	1%	1%	2%	0%	0%	1%	0%	0%	1%	0%	0%	0%
FR	0%	0%	0%	1%	3%	**6%**	0%	0%	0%	0%	0%	0%
GR	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
HU	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
IE	2%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
IS	0%	0%	0%	1%	1%	0%	0%	0%	0%	0%	0%	0%
IT	0%	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%	2%
LT	0%	0%	0%	0%	0%	0%	**11%**	**7%**	**13%**	1%	1%	2%
LU	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
LV	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
MT	0%	0%	0%	0%	1%	3%	0%	0%	0%	0%	0%	0%
NL	1%	0%	0%	**5%**	2%	2%	0%	0%	0%	0%	0%	0%
NO	2%	0%	0%	**9%**	**10%**	**4%**	0%	0%	0%	0%	0%	0%
PL	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
PT	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
RO	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
SE	0%	0%	**4%**	1%	**4%**	**5%**	0%	0%	0%	0%	0%	0%
SI	0%	0%	0%	3%	1%	3%	0%	0%	0%	0%	0%	0%
SK	0%	1%	0%	0%	0%	0%	0%	0%	0%	1%	1%	1%
UK	**8%**	3%	**5%**	1%	1%	1%	0%	0%	0%	0%	0%	0%
Total	0%	0%	1%	1%	1%	1%	1%	1%	1%	0%	0%	0%

For this purpose, we confine the study to the persons who have most often answered the survey questions on their own behalf. This reduces the missing data for all the other variables in the individual file. But significant proportions of missing data remain: for the "Education" variable to a large extent in Denmark and to a lesser degree in Sweden among the over-80s, and in the United Kingdom among the 30-69s and over-79s; for the "Occupation" variable in France, the Netherlands, Norway, and Sweden; and for the "Self-perceived health" variable in Lithuania and Estonia.

III. Distortion of sample and adjustment of prevalence of activity limitations

In this section, we examine the distortion between the final sample and the weighted EU-SILC database (which is supposed to represent the distribution closest to reality) caused by the selection

of persons aged 30+ who responded to the "activity limitation" question. We take the "age", "highest education level attained", and "occupation" variables. We calculate the absolute differences between the percentages of each education level (or occupation) in the final and weighted EU-SILC samples. The distortions are expressed as the ratio of the sum of these values to the number of categories (for example, three for education levels). We initially concentrate on the variations without weighting the final sample, so as to capture any selection biases for the persons who eventually answer these questions. We then look at the variations with weighted samples. The weighting system should reduce the differences without, however, eliminating them altogether. We conclude with a detailed look at which categories are under- or over-represented.

Table 5: Comparison of raw distributions of final sample and weighted distributions of EU-SILC for "Age", "Occupation" and "Education" variables, by country (mean spreads between proportion in final sample and proportion in EU-SILC)

	30+			30-79		80+	
	Occupation	**Education**	**Age**	**Occupation**	**Education**	**Occupation**	**Education**
AT	0.06‰	0.10‰	0.07‰	0.07‰	0.10‰	0.00‰	0.00‰
BE	0.00‰	0.13‰	0.23‰	0.00‰	0.11‰	0.00‰	0.33‰
BG	0.03‰	0.02‰	0.13‰	0.02‰	0.03‰	0.23‰	0.21‰
CY	0.07‰	0.03‰	0.07‰	0.05‰	0.08‰	0.43‰	0.41‰
CZ	**2.14‰**	**2.61‰**	**3.15‰**	**2.25‰**	**2.76‰**	**3.03‰**	**4.36‰**
DE	0.16‰	0.39‰	0.22‰	0.18‰	0.43‰	0.85‰	0.84‰
DK	**3.77‰**	**7.94‰**	**9.63‰**	**3.80‰**	**5.67‰**	**8.53‰**	**12.07‰**
EE	0.00‰	0.00‰	0.31‰	0.00‰	0.00‰	0.00‰	0.00‰
ES	0.01‰	0.00‰	0.31‰	0.01‰	0.00‰	0.00‰	0.00‰
FI	**2.84‰**	**8.49‰**	**6.18‰**	**2.70‰**	**5.99‰**	**10.43‰**	**5.37‰**
FR	0.06‰	0.09‰	0.16‰	0.07‰	0.04‰	0.26‰	0.73‰
GR	0.00‰	0.00‰	0.27‰	0.00‰	0.00‰	0.00‰	0.00‰
HU	0.44‰	0.71‰	0.34‰	0.43‰	0.70‰	1.48‰	1.83‰
IE	0.02‰	0.05‰	0.02‰	0.02‰	0.05‰	0.00‰	0.00‰
IS	**2.19‰**	**3.29‰**	**8.24‰**	**2.69‰**	**1.70‰**	**7.21‰**	**3.31‰**
IT	0.26‰	0.34‰	0.15‰	0.28‰	0.34‰	0.36‰	0.44‰
LT	**0.98‰**	**1.11‰**	**0.82‰**	**1.01‰**	**1.09‰**	**1.09‰**	**2.15‰**
LU	0.39‰	1.11‰	0.49‰	0.40‰	1.20‰	0.00‰	0.00‰
LV	0.00‰	0.00‰	0.48‰	0.00‰	0.00‰	0.00‰	0.00‰
MT	0.02‰	0.04‰	0.05‰	0.02‰	0.04‰	0.00‰	0.00‰
NL	**2.53‰**	**4.07‰**	**7.15‰**	**2.70‰**	**2.46‰**	**7.13‰**	**8.89‰**
NO	**3.87‰**	**2.02‰**	**7.16‰**	**3.69‰**	**3.20‰**	**5.88‰**	**8.54‰**
PL	0.01‰	0.01‰	1.34‰	0.01‰	0.01‰	0.00‰	0.00‰
PT	0.08‰	0.14‰	0.28‰	0.08‰	0.15‰	0.00‰	0.00‰
RO	0.00‰	0.00‰	0.18‰	0.00‰	0.00‰	0.00‰	0.00‰
SE	**4.78‰**	**4.91‰**	**10.00‰**	**3.99‰**	**2.76‰**	**0.00‰**	**12.51‰**
SI	**3.95‰**	**3.93‰**	**9.27‰**	**3.66‰**	**3.91‰**	**16.32‰**	**32.08‰**
SK	0.24‰	0.48‰	0.44‰	0.26‰	0.47‰	0.97‰	1.18‰
UK	0.11‰	0.20‰	0.33‰	0.12‰	0.23‰	0.00‰	0.00‰

All the countries having conducted telephone interviews are among those exhibiting the greatest distortions owing to persons not contacted directly, for whom a proxy may have provided information on the other parts of the questionnaire (except health). Lithuania, the Czech Republic, and Poland exhibit relatively high distortions. In Poland, the distortions are confined to age

distributions. Logically enough, the distortions are very low in those of the remaining countries that display fairly modest attrition. The distortions are greater at ages 80+.

In all countries with major distortions, we observe an increase in the mean age of the national sample. The economically active ages are more specifically under-represented in countries using CATI, with an under-representation of the 40-60s (a group displaying greater attrition) and an over-representation of the over-64s. For the other countries, the distortions are comparable but generally at younger ages: under-representation starts at 30, over-representation at around 55.

Education levels 0-2 are generally over-represented in the final sample relative to EU-SILC data, particularly in Denmark and Finland. The proportions of the higher levels decrease in the final sample. These variations may be due to a cohort effect, as the cohorts that are more represented in the survey because they are older and available completed their studies in periods where education times were shorter than for the more recent cohorts. “Clerical support workers”, “service and sales workers”, and “elementary occupations” are better represented in the final sample. “Managers”, “professionals”, “technicians and associate professionals”, and “craft and related trades workers” are less represented. The variations in education levels and in occupation categories and ages are linked. The distortions are generally found for all three types of variables and are consistent. They concern education levels 0-2, clerical support workers, service workers, and elementary occupations.

After weighting (Table 6), the strong distortions previously observed are eliminated for all countries with high proportions of missing data for the “Activity limitation” variable. The specific weighting variable for these countries is documented for only a part of the population surveyed, confining the impact of the distortions to the sub-sample of persons selected at random and responding to the interview on an individual basis. The special case of Denmark, which now displays no distortion whatsoever, is due to the fact that all persons with adjustment weightings document the “Activity limitation” variable. For these countries (DK, FI, IS, NL, NO, SE, SI), we thus need to compare the distributions with reference data in order to assess the quality of the final sample.

Figures 3 and 4 show the distributions for the expected population size (reference data) and for the weighted population size in the final sample, then in percentages (with distributions based on the raw population size). By tabulating population size, we can identify potential under-representations of certain groups and so examine the determinants that generated this selection effect. The over-representation of other groups will have a greater influence on the final structure of the sample. The reason is that the health data are potentially affected by these two effects. First, the selection effect may be stronger in certain groups. The determinants of the selection are more or less health-related. Second, the sample distortion effect shifts the mean values towards the values observed in groups that are over-represented with respect to the likely results of the reference distribution. While we cannot explore the characteristics and intensity of the selection effect, we can assess the scope of the distortion-related bias by “adjusting” it. We give the results below. The EU-SILC data shown and compared are weighted.

1. Representativeness of age in final sample

The distributions of respondents by age group—expressed in percentages—are generally consistent with the reference data. This is partly due to the fact that all countries use age as an adjustment variable. The distributions in percentages by country show the over-representation of the older age groups in most countries, and the under-representation of younger age groups, revealed earlier by the age distribution of the total EU-SILC database (Figure 1) and largely adjusted after weighting.

Table 6: Comparison of weighted distributions in final sample and weighted distributions in EU-SILC for "Age", "Occupation" and "Education" variables, by country (mean spreads between proportion in final sample and proportion in EU-SILC)

	30+			30-79		80+	
	Occupation	**Education**	**Age**	**Occupation**	**Education**	**Occupation**	**Education**
AT	0.07‰	0.10‰	0.08‰	0.07‰	0.10‰	0.00‰	0.00‰
BE	0.00‰	0.08‰	0.19‰	0.00‰	0.07‰	0.00‰	0.16‰
BG	0.03‰	0.03‰	0.03‰	0.02‰	0.03‰	0.36‰	0.34‰
CY	0.04‰	0.02‰	0.04‰	0.02‰	0.03‰	0.48‰	0.44‰
CZ	**2.20‰**	**2.20‰**	**3.23‰**	**2.32‰**	**2.49‰**	**4.19‰**	**7.16‰**
DE	0.20‰	**0.53‰**	0.35‰	0.22‰	**0.58‰**	**0.78‰**	**0.62‰**
DK	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰
EE	0.00‰	0.00‰	0.39‰	0.00‰	0.00‰	0.00‰	0.00‰
ES	0.01‰	0.00‰	0.01‰	0.01‰	0.00‰	0.00‰	0.00‰
FI	**0.64‰**	**1.03‰**	0.31‰	**0.61‰**	**0.97‰**	**1.94‰**	**1.16‰**
FR	0.05‰	0.10‰	0.05‰	0.05‰	0.14‰	0.30‰	0.21‰
GR	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰
HU	**0.50‰**	0.46‰	0.41‰	0.49‰	0.41‰	**1.66‰**	**1.79‰**
IE	0.03‰	0.08‰	0.03‰	0.04‰	0.08‰	0.00‰	0.00‰
IS	0.27‰	**0.58‰**	0.24‰	0.29‰	**0.64‰**	0.00‰	0.00‰
IT	0.22‰	0.23‰	0.12‰	0.25‰	0.24‰	0.45‰	0.39‰
LT	**0.99‰**	**1.23‰**	**0.57‰**	**1.02‰**	**1.24‰**	**1.82‰**	**0.62‰**
LU	0.25‰	**0.61‰**	0.35‰	0.26‰	**0.68‰**	0.00‰	0.00‰
LV	0.00‰	0.00‰	0.49‰	0.00‰	0.00‰	0.00‰	0.00‰
MT	0.01‰	0.02‰	0.03‰	0.01‰	0.03‰	0.00‰	0.00‰
NL	0.13‰	0.26‰	0.15‰	0.15‰	0.26‰	0.41‰	**0.60‰**
NO	0.31‰	**0.97‰**	0.39‰	0.37‰	**0.86‰**	**3.40‰**	**1.52‰**
PL	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰
PT	0.10‰	0.19‰	0.11‰	0.11‰	0.20‰	0.00‰	0.00‰
RO	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰	0.00‰
SE	0.05‰	0.03‰	0.09‰	0.05‰	0.04‰	0.00‰	0.00‰
SI	0.06‰	0.06‰	0.06‰	0.06‰	0.06‰	0.00‰	0.00‰
SK	0.29‰	**0.63‰**	0.45‰	0.30‰	**0.65‰**	**1.18‰**	**1.53‰**
UK	0.09‰	0.22‰	0.26‰	0.10‰	0.24‰	0.00‰	0.00‰

IV. Representativeness of sample used to document health variables

We analyze the representativeness of the final sample—which contains the health data examined in this study—using statistical data on the socio-demographic distributions of European populations. We want to measure potential distortions due to a selection effect in the final sample relative to the total EU-SILC database. We use two data sources published by the International Labour Office (ILO) for the distribution across occupations and by Eurostat for the distribution across Educational levels, using information from the Labour Force Survey (EU-LFS). Although we are aware that these data are also dependent on the level of participation of the individuals to the data collection process and knowing European variation in the study design, it is for us an opportunity to look at a certain level of congruence from one source to another.

Figures 3 and 4 show the distributions for the expected population size (reference data) and for the weighted population size in the final sample, then in percentages (with distributions based on the

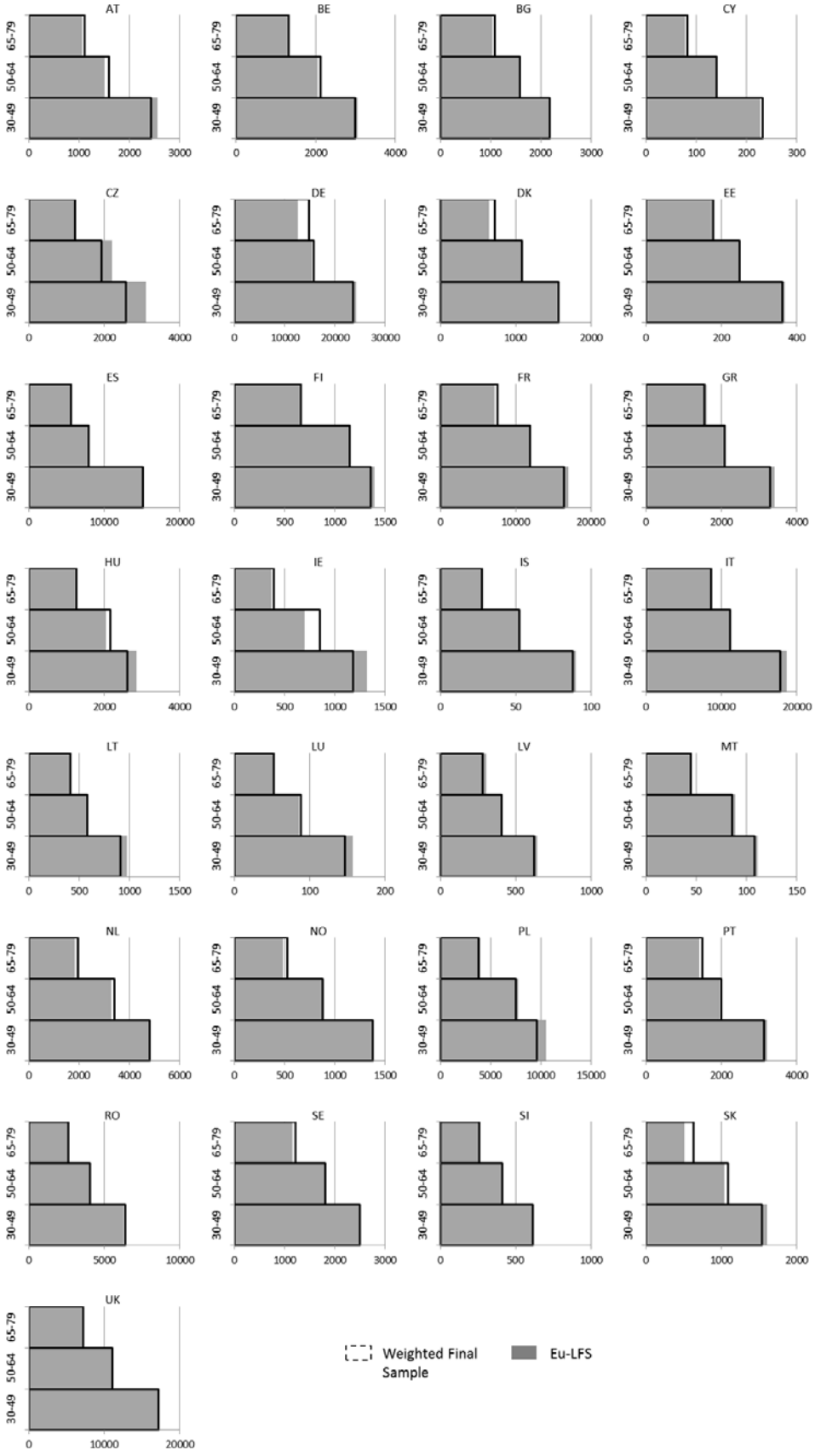
raw population size). By tabulating population size, we can identify potential under-representations of certain groups and so examine the determinants that generated this selection effect. The over-representation of other groups will have a greater influence on the final structure of the sample. The reason is that the health data are potentially affected by these two effects. First, the selection effect may be stronger in certain groups. The determinants of the selection are more or less health-related. Second, the sample distortion effect shifts the mean values towards the values observed in groups that are over-represented with respect to the likely results of the reference distribution. While we cannot explore the characteristics and intensity of the selection effect, we can assess the scope of the distortion-related bias by "adjusting" it. We give the results below. The EU-SILC data shown and compared are weighted.

1. Representativeness of age in final sample

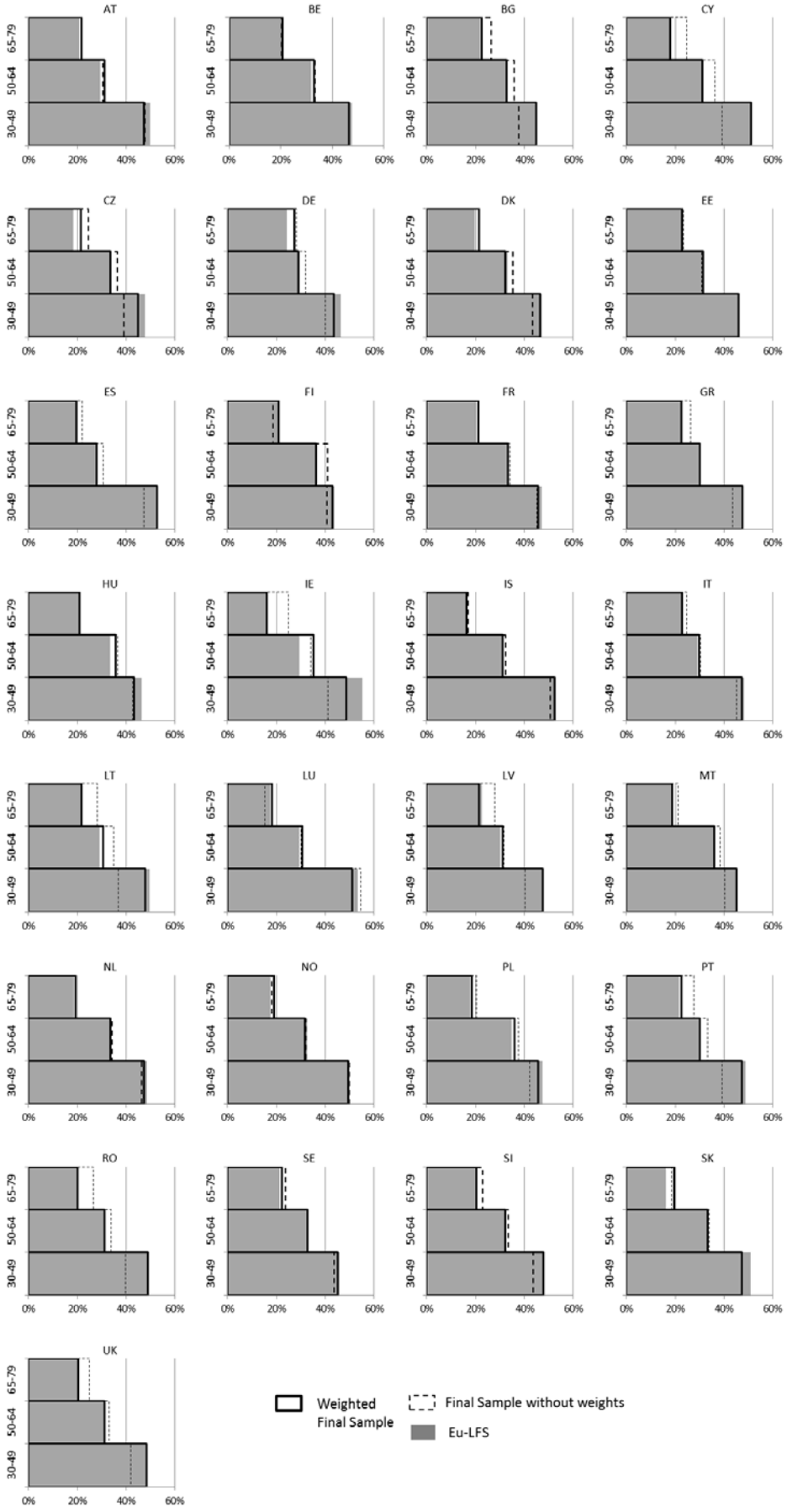
The distributions of respondents by age group—expressed in percentages—are generally consistent with the reference data. This is partly due to the fact that all countries use age as an adjustment variable. The distributions in percentages by country show the over-representation of the older age groups in most countries, and the under-representation of younger age groups, revealed earlier by the age distribution of the total EU-SILC database (Figure 1) and largely adjusted after weighting.

Figure 2: Expected distributions (EU-LFS) and observed distributions (weighted final sample) of age groups

a. By population size



b. In percentages



2. Representativeness of "Education" variable in final sample

We can also analyze distributions of respondents by education level using data from the final sample and Eurostat. The latter are taken from the EU-LFS (Labour Force Survey) and can serve as reference data because they are closely aligned with national data and because the EU-LFS covers a wider base of respondents[5]. As the Eurostat data are available for five-year age groups, the variations are derived from standardized frequencies. The distributions of education levels by country exhibit absolute variations in frequencies of less than 10%, except in Luxembourg and Malta for aggregated levels 0-2 (+12% and -12% respectively). Recall that few countries used education level as an adjustment variable for the weightings. Accordingly, we find weighted EU-SILC distributions calibrated on the reference distributions in these countries, but discrepancies persist in many cases. The weighted EU-SILC data yield a number of respondents very close to that provided by the reference data for 14 countries (AT, BG, DE, CZ, DK, FI, HU, LV, IT, NO, PL, PT, RO, SI). In terms of population size, the weights of high-attrition countries enable us to make up for the losses.

In other words, in the distributions, the low education levels (0-2) tend to be mildly under-represented (CY, ES, FR, NL, SK) or more significantly so (GR, MT, SE, UK, IS) in most countries. In these countries, the sample is larger than expected at levels 3-4 (FR, MT, NL, UK, SE, IS) and/or levels 5-6 (GR, NL, MT, SK, SE). By contrast, low education levels are over-represented in four countries (IE, LU, LT, EE). Belgium has the expected number of respondents for levels 0-2 but over-represents the highest level; the intermediate group is smaller than expected. Luxembourg and Iceland are special cases, since their distributions diverge substantially from their expected values.

The percentage charts also show that, for many countries, weighting has little impact on the distributions of respondents across the three education groups. Most of the variations that do appear consist of under-representations of levels 0-2 offset by over-representations of levels 5-6. This is notably the case for DE, DK, NO, and—to a lesser extent—FR, IE, NL, and IT (with an over-representation of 3-4 for IT only). We can connect these variations with the observed under-representations of the unweighted 30-49 age groups, generally associated with higher education levels. CY and EE display other variations, with slight over-representations of levels 0-2 and under-representations of high levels.

These variations should also be viewed against the national weighting methods. Some countries (notably Denmark, France, and Slovenia) have treated education as an adjustment variable for their national reference database in determining the weightings for individuals. These countries actually exhibit fairly small mean variations, even in the samples used to document the health variables. Not all the countries follow this pattern. In others, the distributions of the EU-SILC database and the final sample diverge somewhat from the reference distribution.

For Luxembourg, Malta, and Iceland, the final sample under-represents or over-represents one of the three education groups by more than ten percentage points. These under- or over-representations are not adjusted by the weighting methods. It seems hard to use these samples to study disability levels in these groups and their differences. Countries with under-representation of the lower educated are: SE, CY, GR, ES, FR, HU, NL, SK, UK. Countries with over-representation of the lower educated are: EE, LT, IE. Regarding the higher educated, the under and over-representations are less important. BG, EE have an under representation and BE, ES, GR, SE, SK.

As we lack suitable data to determine whether these selection effects are health-related, we can only make assumptions. Poor health is one of the factors causing non-response. However, it is somewhat offset by non-response from people who lack the time to participate and are in relatively good health. In other words, under these assumptions, respondents and non-respondents differ in regard to their health-risk exposure. As a result, the weighting systems cannot always adjust health

[5] S Schneider and W. Müller, "Measurement of Education in EU-SILC – Preliminary Evaluation of Measurement Quality", Equasolc Working Paper 2009/5. http://www.equalsoc.org/publication/283 (retrieved 21 November 2011).

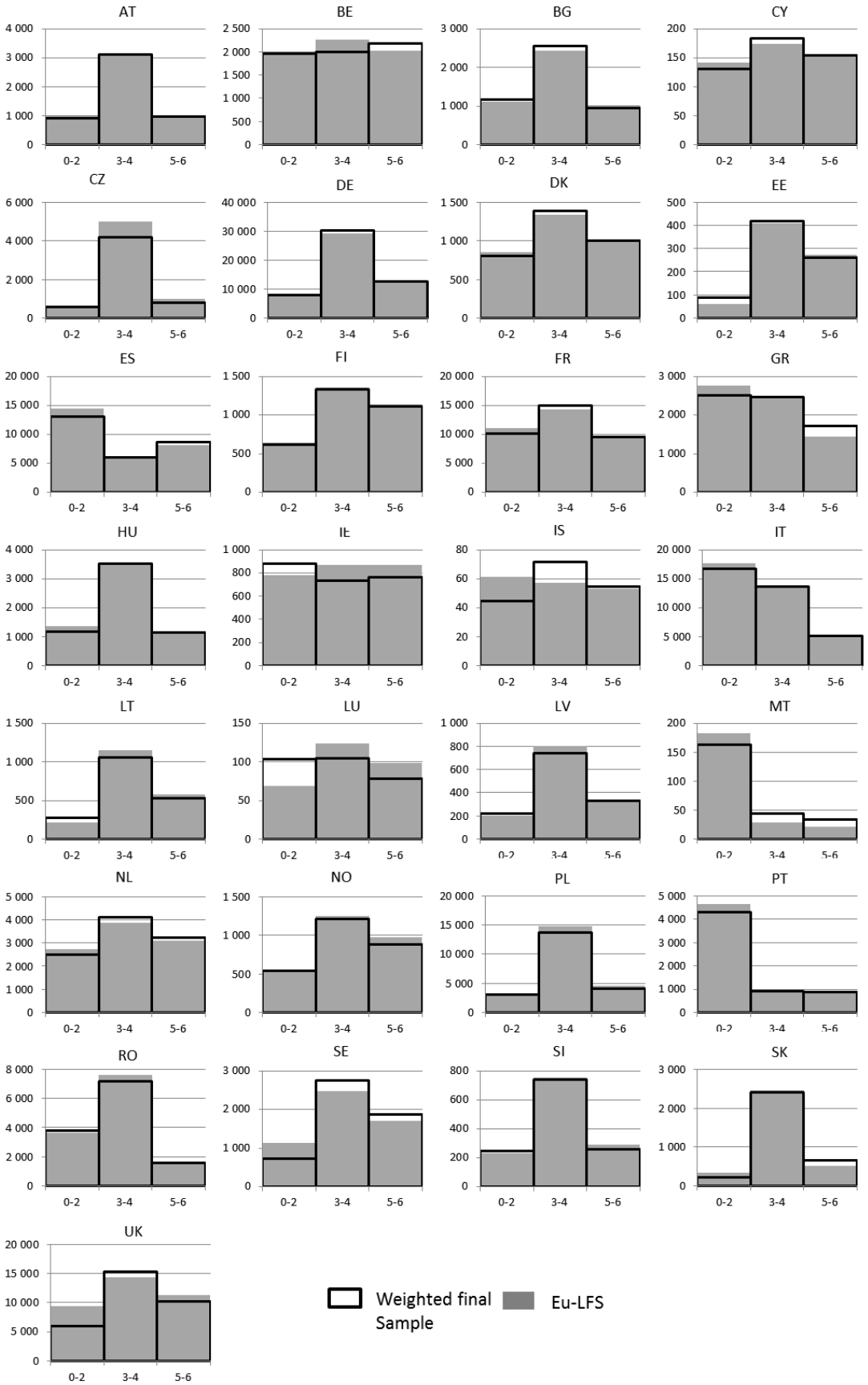
estimates correctly, since they are based on “similar individuals" but with different health risks. A study of Belgian data effectively showed that non-respondents to the national health survey (whose sample is linked to the census, which includes health questions) tend to belong to socially disadvantaged groups and are in poor health. The authors argue that the activity limitations in these groups are therefore under-estimated and the gaps with the other groups are understated (Lorant *et al*. 2007). The analysis of the EU-SILC samples shows that, in a number of countries, the selection effect of was stronger in the less educated groups, consistently with the previous finding. Yet we also find a selection effect in the highest-educated groups as well as among the youngest age groups and economically active persons compared with the economically inactive. As mentioned earlier, being unavailable to respond is a cause of non-response; by contrast, being at home, including for health reasons, can facilitate response.

Table 8: Distributions of highest education levels attained by country and absolute variations from Eurostat reference distributions in EU-SILC, for persons aged 25-69

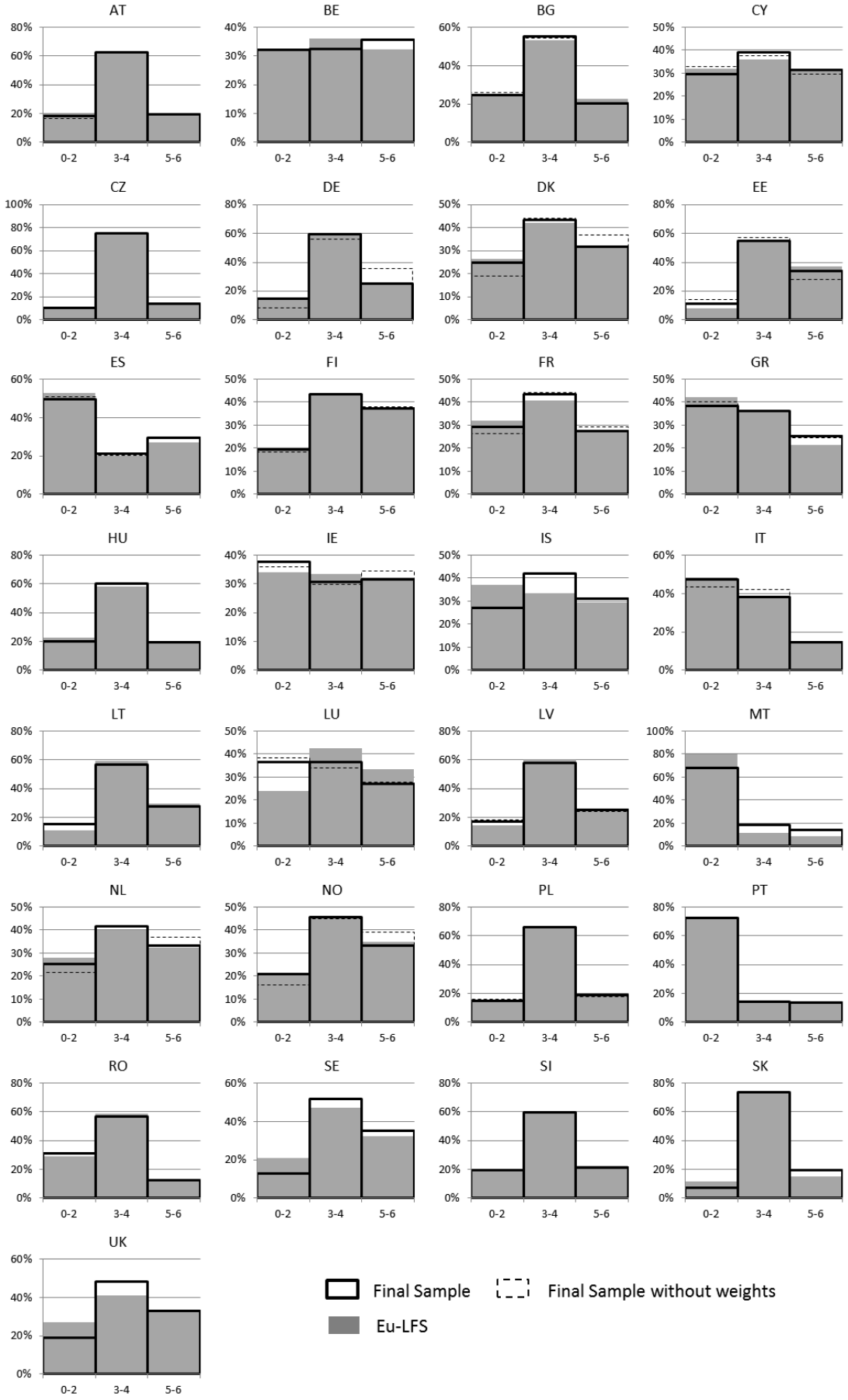
	0-2	3-4	5-6	Mean |Variation|
AT	18.3% (-1.3)	62.4% (0.6)	19.3% (0.7)	0.86%
BE	32.0% (0.1)	32.4% (-3.6)	35.6% (3.5)	2.37%
BG	24.8% (0.5)	55.0% (1.9)	20.2% (-2.4)	1.60%
CY	29.6% (-2.4)	39.0% (2.1)	31.4% (0.4)	1.61%
CZ	10.5% (0.9)	75.2% (-0.1)	14.3% (-0.8)	0.60%
DE	15.1% (-0.2)	59.8% (1.0)	25.2% (-0.9)	0.70%
DK	24.8% (-1.7)	43.5% (1.5)	31.6% (0.2)	1.10%
EE	11.2% (3.3)	54.9% (-0.2)	33.9% (-3.1)	2.20%
ES	49.5% (-3.2)	21.1% (1.0)	29.4% (2.3)	2.16%
FI	19.3% (-0.4)	43.6% (0.2)	37.2% (0.2)	0.27%
FR	29.1% (-2.7)	43.4% (2.6)	27.5% (0.1)	1.80%
GR	38.3% (-4.0)	36.3% (-0.1)	25.4% (4.0)	2.69%
HU	19.9% (-2.8)	60.4% (2.3)	19.7% (0.5)	1.86%
IE	37.8% (3.6)	30.6% (-2.9)	31.6% (-0.7)	2.42%
IS	**26.9% (-10.0)**	**41.9% (8.3)**	**31.2% (1.7)**	**6.70%**
IT	47.2% (-1.3)	38.3% (0.6)	14.4% (0.7)	0.88%
LT	15.4% (4.2)	56.8% (-2.4)	27.8% (-1.7)	2.77%
LU	**36.6% (12.5)**	**36.3% (-6.3)**	**27.1% (-6.3)**	**8.36%**
LV	17.0% (2.2)	57.8% (-2.4)	25.2% (0.1)	1.58%
MT	**67.8% (-12.0)**	**18.4% (6.7)**	**13.8% (5.3)**	**8.00%**
NL	25.2% (-2.7)	41.7% (1.8)	33.1% (0.9)	1.82%
NO	20.8% (0.5)	45.8% (0.9)	33.4% (-1.4)	0.91%
PL	15.0% (0.3)	65.9% (0.2)	19.1% (-0.6)	0.37%
PT	72.2% (-0.6)	14.0% (0.4)	13.7% (0.3)	0.42%
RO	31.0% (1.7)	56.8% (-1.8)	12.2% (0.1)	1.18%
SE	13.0% (-7.7)	51.7% (4.7)	35.3% (3.0)	5.14%
SI	19.5% (1.1)	59.6% (0.6)	20.9% (-1.7)	1.12%
SK	6.9% (-4.5)	73.4% (-0.3)	19.6% (4.8)	3.17%
UK	19.0% (-8.0)	48.3% (7.5)	32.7% (0.4)	5.32%

Figure 3: Expected distributions (EU-LFS) and observed distributions (weighted final sample) of education levels by population size and in percentages. The percentage charts display unweighted and weighted distributions.

a. By population size



b. In percentage



3. Representativeness of “Occupation” variable in final sample

To refine the description, we can also compare the variations in distribution by socio-occupational category between EU-SILC data and the reference data. The data for the economically active come from the ILO. People are classified as economically active if they are:

- in employment: paid employees, the self-employed, and employers, aged 15+, who have worked even one hour in a reference week or have been temporarily absent because of illness (of under a year’s duration) or leave
- unemployed: persons aged 15+, who have not worked even one hour during a reference week, are available for work in fifteen days or less, and have actively sought a job in the previous month or have found one starting in less than three months.

The “Occupation” variable indicates the type of occupation for the most recent job held. To match the ILO definition with the data for the final sample, we selected respondents as follows: the persons comparable with ILO data are those who document the “Occupation” variable, are aged 15+, and define themselves as paid employees or self-employed. The comparisons are extended to persons aged 15+ who document the “Occupation" variable and describe themselves as actively seeking work and available for work—two criteria that characterize the unemployed. The small number of responses to the variable “has worked at least one hour in the previous week” prevented us from refining these categories. Table 9 shows the absolute variations obtained and highlights the groups subject to the strongest variations.

The results are fairly homogeneous. Most groups do not differ by more than 5% between the weighted distributions of the final sample and the ILO data, except “technicians and associate professionals” in Germany (+5.4%) and “skilled agricultural, forestry and fishery workers” in Portugal (-6.2%).

The variations in distributions observed in Luxembourg are significant but should be viewed in perspective, for they are being compared with the far earlier data from the 2001 Luxembourg census. This is consistent, however, with the observation regarding the "Education" variable, i.e., an over-representation of elementary occupations in manufacturing and services (generally low education level) and an under-representation of management occupations (generally high education level). We also note the absence of agricultural occupations and heavy-vehicle drivers. Similarly, Iceland’s profile for the “Education” variable (under-representation of low education levels) is consistent with its under-representation of unskilled occupations and “technicians and associate professionals” versus an over-representation of (higher-level) professionals. In the United Kingdom, the under-representation of lesser-educated groups goes hand in hand with a larger than expected proportion of professionals in intellectual and administrative fields, and service workers, but a smaller than expected share of industrial occupations.

Professionals are over-represented in several countries (IS, NO, NL, SK), as are “technicians and associate professionals” (DE, CY, BE, NL, SI, SK), and “clerical support workers” in most countries, to a moderate extent but more significantly in Spain.

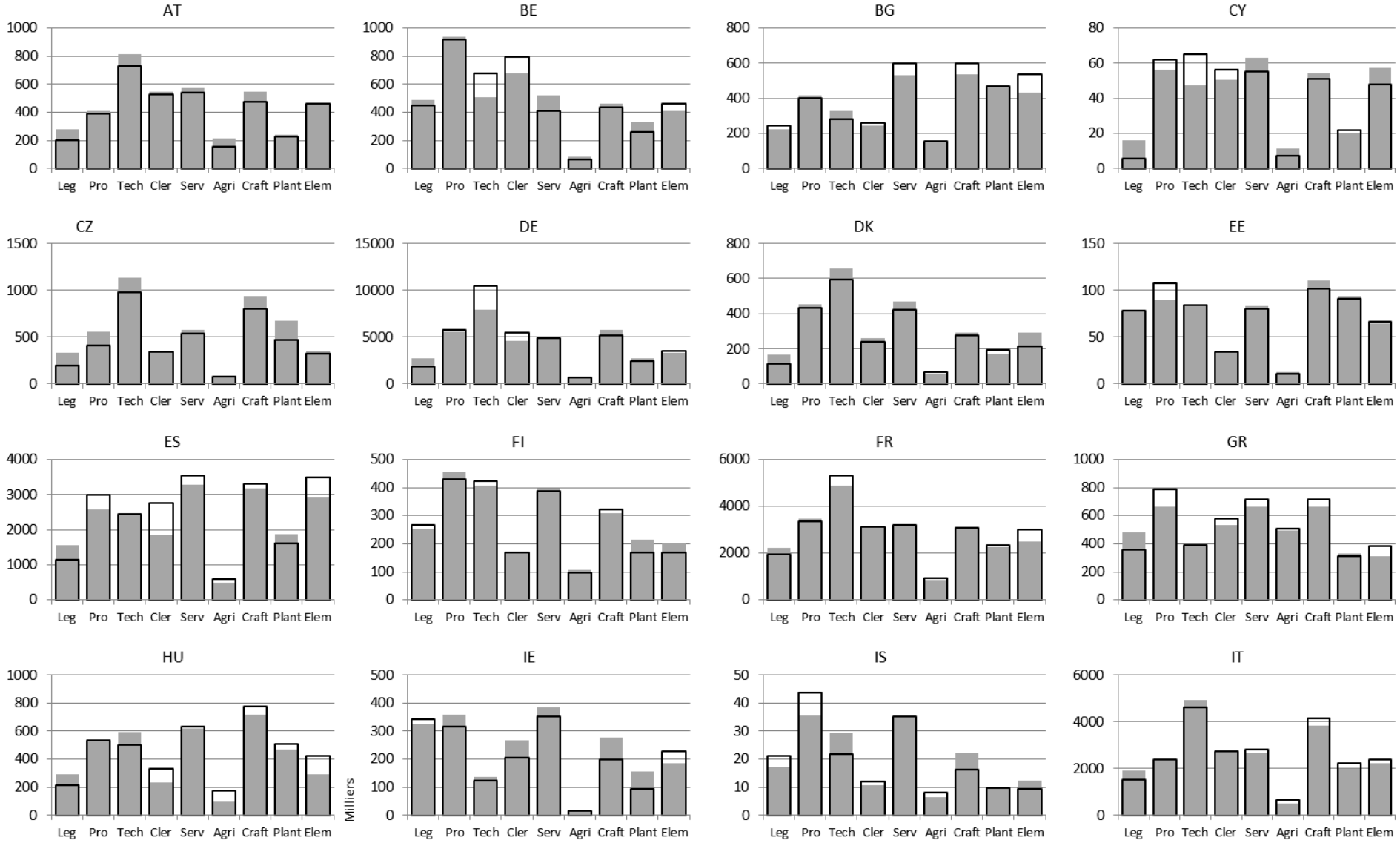
The observed variations in distributions of occupational categories and education levels may also be related to an age effect. The levels generally under-represented are correlated with older ages. For occupational categories, the mean ages are close to 45; for education levels, they are close to 50. The over-represented categories are correlated with younger mean ages.

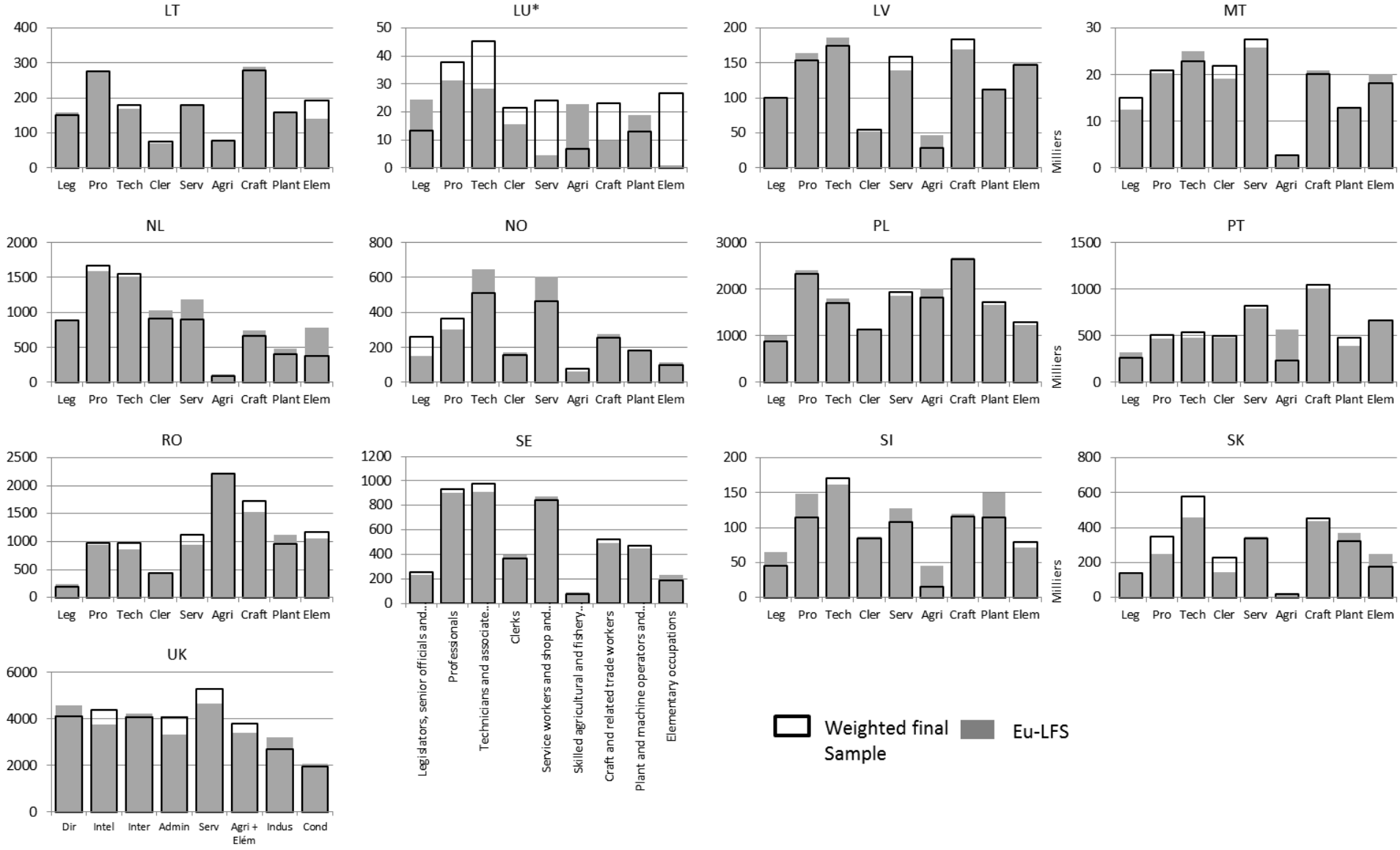
Table 9: Distributions of occupations by country and absolute variations relative to ILO reference distributions in EU-SILC[6]

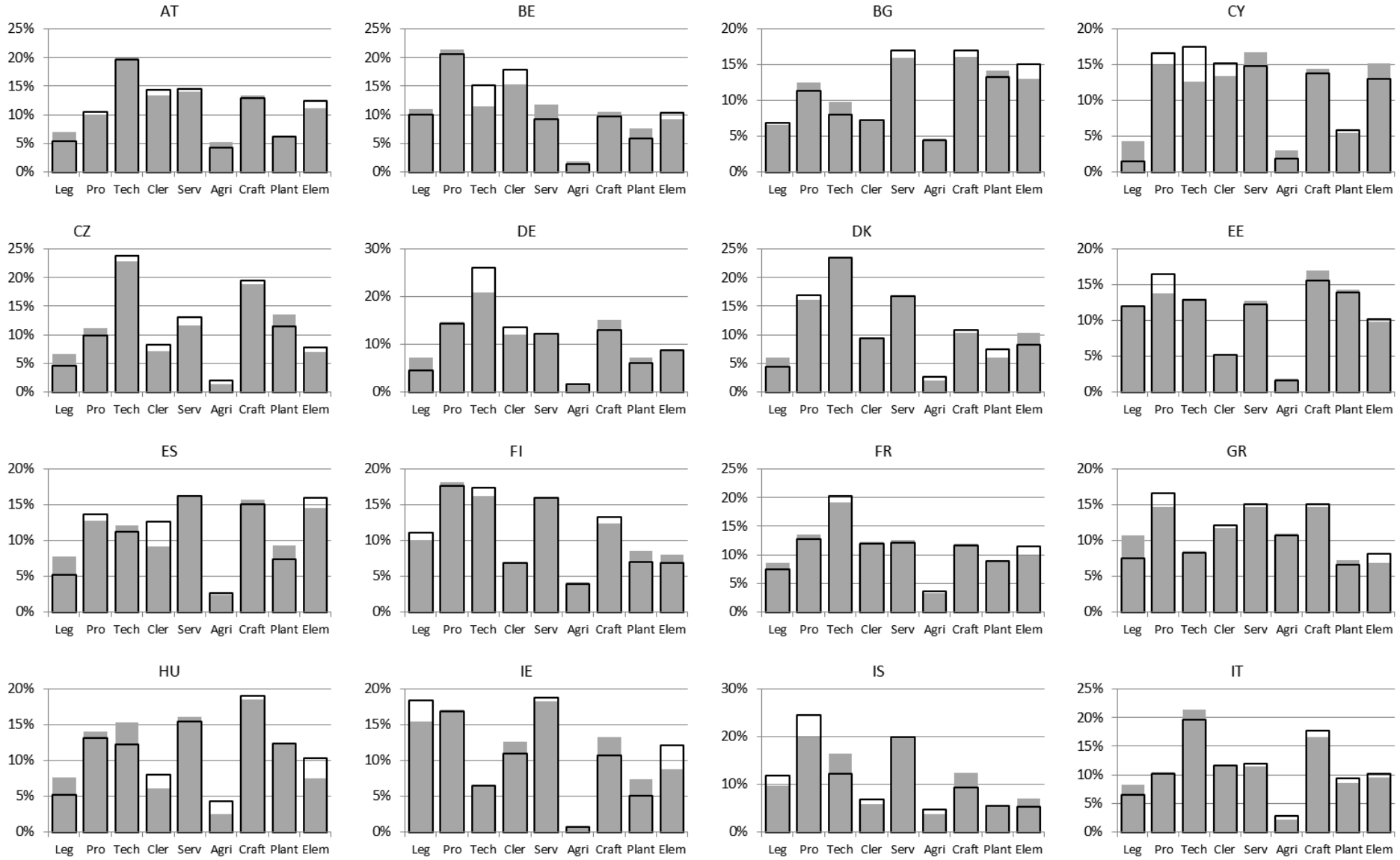
	Armed forces	Legislators, senior officials, and managers	Professionals	Technicians and associate professionals	Clerks	Service workers and shop and market sales workers	Skilled agricultural and fishery workers	Craft and related trades workers	Plant and machine operators, and assemblers	Elementary occupations	Mean |Variation|
AT	0.3% (-0.1)	5.4% (-1.5)	10.4% (0.5)	19.6% (-0.3)	14.2% (0.9)	14.5% (0.5)	4.2% (-1.0)	12.8% (-0.5)	6.1% (0.2)	12.4% (1.3)	0.68%
BE	0.0% (-0.8)	10.0% (-1.0)	20.5% (-0.6)	15.2% (3.9)	17.8% (2.6)	9.1% (-2.5)	1.4% (-0.4)	9.7% (-0.7)	5.8% (-1.7)	10.3% (1.1)	1.52%
BG	1.1% (1.1)	6.7% (0.1)	11.2% (-1.2)	7.9% (-1.9)	7.2% (-0.1)	16.7% (0.7)	4.4% (-0.3)	16.7% (0.7)	13.1% (-1.0)	14.9% (1.8)	0.90%
CY	1.5% (0.2)	1.5% (-2.8)	16.4% (1.6)	17.2% (4.8)	14.9% (1.6)	14.6% (-1.9)	1.9% (-1.1)	13.6% (-0.6)	5.7% (0.4)	12.7% (-2.2)	1.72%
CZ	0.4% (0.0)	4.6% (-2.0)	9.8% (-1.3)	23.7% (1.0)	8.2% (1.1)	13.0% (1.4)	1.9% (0.6)	19.4% (0.7)	11.4% (-2.2)	7.7% (0.8)	1.11%
DE	0.0% (-0.6)	4.6% (-2.7)	14.4% (-0.2)	26.0% (5.4)	13.5% (1.6)	12.2% (-0.2)	1.6% (-0.2)	12.9% (-2.1)	6.0% (-1.2)	8.8% (0.1)	1.42%
DK	0.5% (0.0)	4.4% (-1.5)	16.8% (0.8)	23.3% (0.1)	9.3% (0.1)	16.6% (0.1)	2.7% (0.7)	10.7% (0.4)	7.5% (1.5)	8.2% (-2.1)	0.73%
EE	0.5% (0.5)	11.9% (-0.3)	16.4% (2.6)	12.8% (-0.2)	5.2% (-0.2)	12.2% (-0.5)	1.6% (-0.2)	15.5% (-1.5)	13.8% (-0.5)	10.1% (0.3)	0.68%
ES	0.6% (0.1)	5.2% (-2.5)	13.6% (0.9)	11.1% (-0.9)	12.5% (3.4)	16.1% (0.0)	2.7% (0.2)	15.0% (-0.7)	7.3% (-1.9)	15.9% (1.4)	1.21%
FI	0.3% (-0.9)	11.0% (1.0)	17.6% (-0.4)	17.3% (1.3)	6.9% (0.3)	15.9% (0.2)	4.0% (-0.2)	13.2% (1.1)	7.0% (-1.4)	6.9% (-1.0)	0.78%
FR	1.2% (0.0)	7.4% (-1.2)	12.6% (-0.8)	20.0% (1.1)	11.8% (-0.2)	12.0% (-0.5)	3.5% (0.3)	11.5% (-0.2)	8.8% (0.0)	11.2% (1.6)	0.60%
GR	1.3% (1.3)	7.4% (-3.2)	16.3% (1.6)	8.1% (-0.4)	12.0% (0.2)	14.9% (0.3)	10.6% (-0.4)	14.9% (0.3)	6.5% (-0.8)	8.0% (1.1)	0.97%
HU	1.1% (0.2)	5.1% (-2.4)	12.9% (-1.0)	12.1% (-3.1)	8.0% (1.9)	15.3% (-0.6)	4.3% (1.8)	18.8% (0.4)	12.2% (0.1)	10.2% (2.7)	1.42%
IE	0.4% (0.1)	18.2% (2.9)	16.7% (-0.3)	6.5% (0.1)	10.9% (-1.7)	18.7% (0.5)	0.8% (0.0)	10.6% (-2.6)	5.1% (-2.3)	12.0% (3.3)	1.39%
IS	0.0% (0.0)	11.9% (2.2)	24.6% (4.6)	12.3% (-4.2)	6.8% (0.9)	19.9% (0.1)	4.6% (0.9)	9.3% (-3.2)	5.5% (0.3)	5.3% (-1.7)	1.81%
IT	0.9% (-0.2)	6.5% (-1.7)	10.0% (-0.3)	19.5% (-1.6)	11.6% (0.1)	11.8% (0.5)	2.8% (0.7)	17.5% (1.2)	9.3% (0.9)	10.0% (0.6)	0.77%
LT	0.2% (0.0)	9.6% (-0.9)	17.7% (-0.4)	11.4% (0.2)	4.7% (0.2)	11.4% (-0.6)	4.9% (-0.3)	17.8% (-1.2)	10.1% (-0.1)	12.2% (3.1)	0.71%
LU	0.4% (-6.5)	6.3% (-8.3)	17.8% (-0.9)	21.4% (4.5)	10.1% (0.8)	11.3% (8.6)	3.2% (-10.4)	10.9% (5.2)	6.1% (-5.1)	12.6% (12.1)	6.24%
LV	0.1% (-0.1)	9.0% (0.0)	13.8% (-0.8)	15.6% (-1.0)	4.9% (0.3)	14.3% (1.9)	2.6% (-1.5)	16.4% (1.3)	10.1% (0.2)	13.2% (-0.2)	0.73%
MT	1.0% (-0.1)	9.1% (1.4)	12.8% (0.1)	14.0% (-1.5)	13.4% (1.5)	16.9% (0.9)	1.6% (-0.1)	12.2% (-0.8)	7.9% (-0.1)	11.1% (-1.3)	0.78%
NL	0.5% (0.1)	11.8% (1.2)	22.2% (3.2)	20.7% (2.6)	12.2% (-0.1)	12.0% (-2.3)	1.2% (-0.1)	8.8% (-0.1)	5.4% (-0.3)	5.1% (-4.2)	1.40%
NO	0.2% (-0.5)	11.0% (5.0)	15.4% (3.5)	21.5% (-4.1)	6.6% (-0.2)	19.6% (-4.3)	3.1% (0.8)	10.7% (-0.2)	7.7% (0.3)	4.2% (-0.2)	1.92%
PL	0.4% (0.0)	5.7% (-0.6)	15.0% (-0.2)	11.0% (-0.3)	7.2% (0.2)	12.5% (0.8)	11.7% (-1.0)	17.0% (0.1)	11.1% (0.6)	8.3% (0.5)	0.43%
PT	0.4% (-0.2)	5.2% (-1.0)	10.0% (1.0)	10.6% (1.4)	9.8% (0.5)	16.2% (1.0)	4.7% (-6.2)	20.7% (1.4)	9.4% (1.9)	13.1% (0.3)	1.49%
RO	0.6% (0.6)	1.9% (-0.5)	9.9% (-0.3)	10.0% (0.8)	4.4% (-0.4)	11.4% (1.3)	22.5% (-1.2)	17.5% (1.2)	9.8% (-2.2)	11.9% (0.7)	0.93%
SE	0.3% (0.1)	5.4% (0.3)	20.1% (0.4)	21.1% (1.3)	7.9% (-0.7)	18.2% (-0.8)	1.6% (-0.4)	11.3% (0.6)	10.1% (0.3)	4.0% (-1.1)	0.61%
SI	0.7% (0.1)	5.3% (-1.4)	13.5% (-1.6)	20.0% (3.5)	9.8% (1.0)	12.7% (-0.3)	1.8% (-2.9)	13.6% (1.4)	13.5% (-1.7)	9.2% (2.0)	1.60%
SK	0.4% (-0.2)	5.3% (-0.2)	13.3% (3.0)	22.1% (3.3)	8.9% (2.8)	12.9% (-1.4)	0.7% (-0.3)	17.4% (-0.6)	12.3% (-3.0)	6.8% (-3.5)	1.83%
UK	0.0% (-0.6)	13.5% (-2.0)	14.5% (1.6)	13.4% (-0.9)	13.4% (2.1)	17.4% (1.6)	12.5% (0.9)	8.8% (-2.1)	6.5% (-0.6)		1.37%

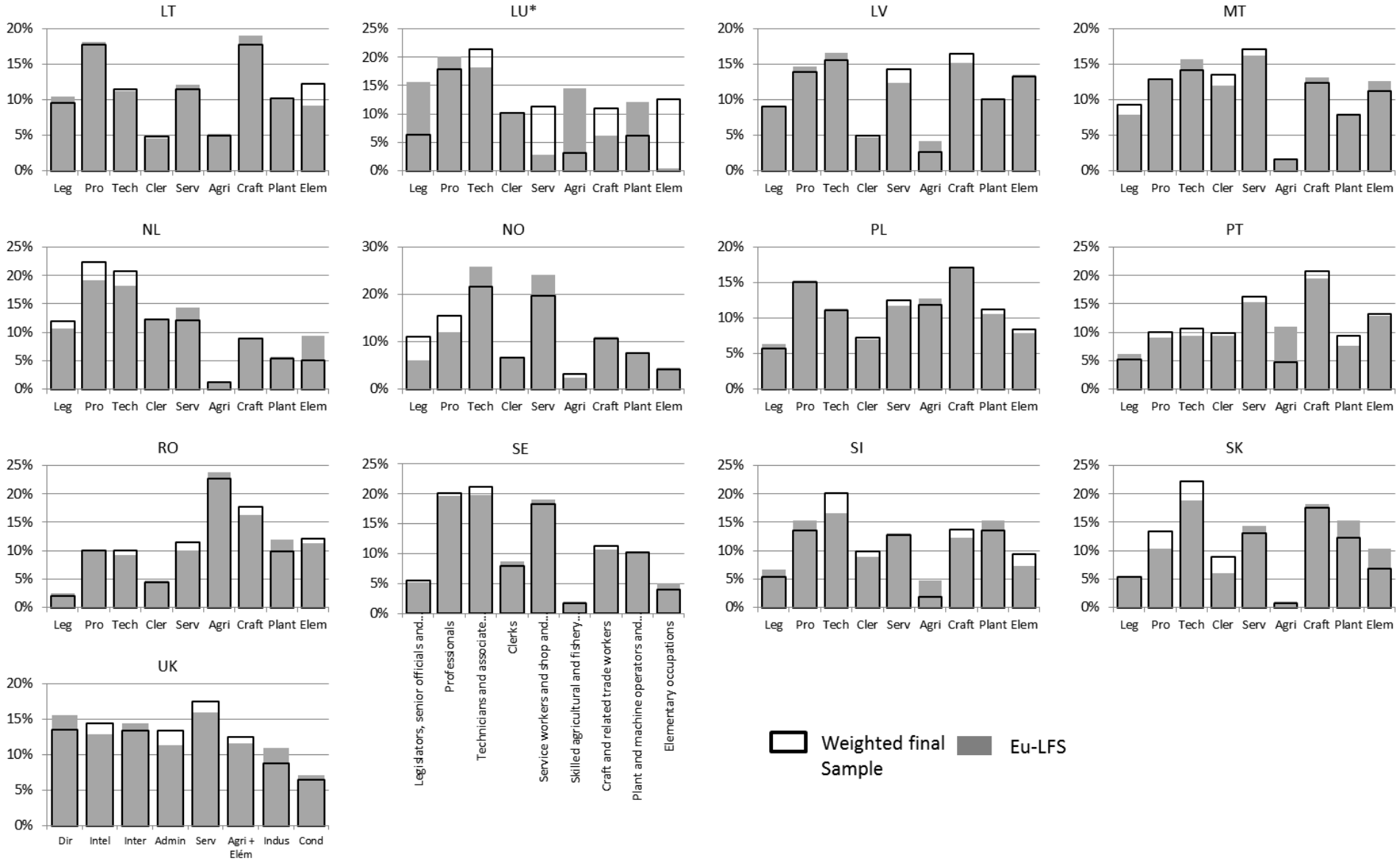
[6] ILO statistics for 2008 for all countries except Luxembourg (national data from 2001 census). For the United Kingdom, the data for "Skilled agricultural, forestry and fishery workers" also include "Elementary occupations".

Figure 4: Expected distributions (ILO) and observed distributions (final sample), weighted and unweighted, of socio-occupational categories (thousands of respondents and percentages).




LT
LU*
LV
MT
NL
NO
PL
PT
RO
SE
SI
SK
UK
Leg Pro Tech Cler Serv Agri Craft Plant Elem
Dir Intel Inter Admin Serv Agri + Elém Indus Cond
Legislators, senior officials and...
Professionals
Technicians and associate...
Clerks
Service workers and shop and...
Skilled agricultural and fishery...
Craft and related trade workers
Plant and machine operators and...
Elementary occupations
Milliers
Weighted final Sample
Eu-LFS


AT
BE
BG
CY
CZ
DE
DK
EE
ES
FI
FR
GR
HU
IE
IS
IT
Leg Pro Tech Cler Serv Agri Craft Plant Elem


LT
LU*
LV
MT
NL
NO
PL
PT
RO
SE
SI
SK
UK
Leg Pro Tech Cler Serv Agri Craft Plant Elem
Dir Intel Inter Admin Serv Agri + Elém Indus Cond
Legislators, senior officials and..
Professionals
Technicians and associate..
Clerks
Service workers and shop and..
Skilled agricultural and fishery..
Craft and related trade workers
Plant and machine operators and..
Elementary occupations
Weighted final Sample
Eu-LFS

V. Prevalence of activity limitations in European Union countries

Table 10 and the charts below give the prevalence of activity limitations by education level. Prevalence is standardized by age for all countries combined. In the countries that distinguish the category of those who report they never had education from low education levels, prevalence is compiled (the number of persons concerned is small). This initial overview yields three conclusions: (1) population distributions by education level vary considerably from one country to another; (2) the prevalence of activity limitations follows the expected gradient in all countries; the "no education" group is very small but displays a very high prevalence of activity limitations; (3) activity limitation levels and gaps vary from one country to another.

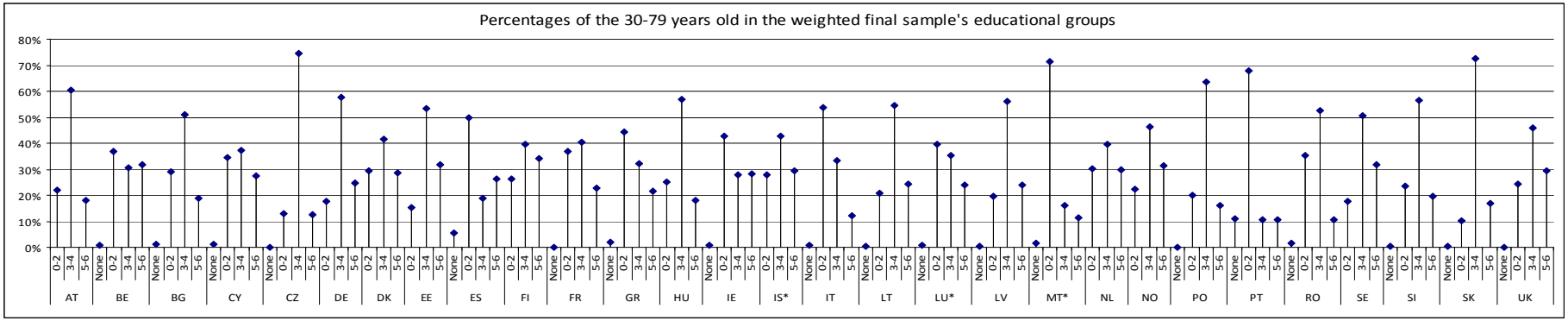
Table 10: Disability levels observed in each country by education level, standard error, population size, and weighted distribution of education levels for the 30-80 age group

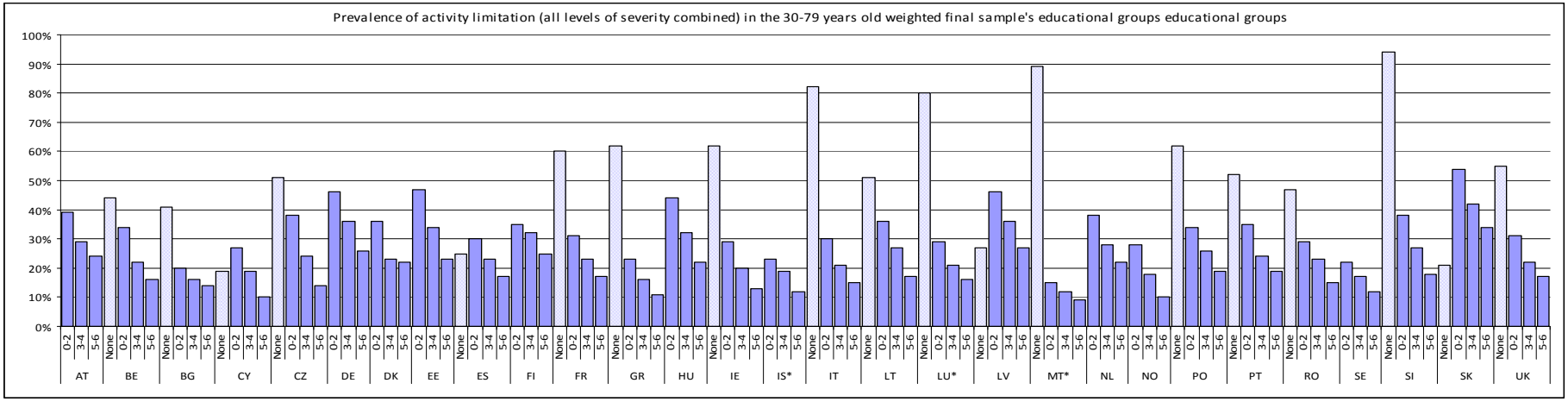
		Raw pop. size	% weighted	Activity limitation			Severe activity limitation		
AT	**0-2**	1750	21.8%	30.2%	39%	[37.1%-41.6%]	10.0%	16%	[14.1%-17.5%]
	3-4	5170	60.2%		29%	[27.5%-30.0%]		9%	[8.5%-10.1%]
	5-6	1579	18.0%		24%	[21.5%-25.6%]		6%	[4.5%-6.8%]
BE	**Never**	44	0.6%	25.0%	44%	[29.6%-59.0%]	8.0%	21%	[9.3%-33.6%]
	0-2	3279	37.0%		34%	[32.4%-35.6%]		12%	[11.1%-13.3%]
	3-4	2670	30.5%		22%	[20.9%-24.1%]		7%	[6.1%-8.0%]
	5-6	2847	31.9%		16%	[15.0%-17.7%]		4%	[3.7%-5.2%]
BG	**Never**	118	1.1%	17.2%	41%	[31.9%-49.7%]	4.4%	30%	[21.4%-37.9%]
	0-2	3195	29.2%		20%	[18.9%-21.7%]		5%	[4.5%-6.1%]
	3-4	4914	50.8%		16%	[15.2%-17.3%]		4%	[3.5%-4.6%]
	5-6	1801	18.9%		14%	[12.3%-15.5%]		4%	[2.8%-4.6%]
CY	**Never**	66	1.0%	21.5%	19%	[9.6%-28.5%]	7.2%	19%	[9.6%-28.5%]
	0-2	2206	34.4%		27%	[25.6%-29.3%]		10%	[9.0%-11.5%]
	3-4	1804	37.1%		19%	[16.7%-20.3%]		5%	[4.3%-6.4%]
	5-6	1256	27.5%		10%	[8.2%-11.5%]		3%	[1.8%-3.5%]
CZ	**Never**	2	0.0%	24.7%	51%	[-18.4%-120%]	6.2%	18%	[-35.2%-71.1%]
	0-2	1856	12.9%		38%	[35.7%-40.1%]		11%	[10.0%-12.9%]
	3-4	10011	74.4%		24%	[23.3%-25.0%]		6%	[5.4%-6.4%]
	5-6	1549	12.6%		14%	[12.6%-16.1%]		3%	[1.8%-3.3%]
DE	**0-2**	2008	17.7%	34.8%	46%	[43.6%-47.9%]	10.9%	18%	[16.3%-19.7%]
	3-4	10584	57.6%		36%	[34.9%-36.7%]		11%	[10.3%-11.5%]
	5-6	6984	24.7%		26%	[25.0%-27.1%]		6%	[5.5%-6.6%]
DK	**0-2**	1112	29.6%	26.2%	36%	[33.2%-38.8%]	8.1%	12%	[9.7%-13.5%]
	3-4	2011	41.7%		23%	[21.1%-24.8%]		7%	[6.1%-8.3%]
	5-6	1579	28.7%		22%	[19.6%-23.7%]		5%	[4.2%-6.4%]
EE	**0-2**	1368	15.1%	31.5%	47%	[44.7%-50.0%]	7.6%	14%	[12.1%-15.7%]
	3-4	4288	53.3%		34%	[32.2%-35.1%]		8%	[7.0%-8.6%]
	5-6	2080	31.6%		23%	[20.8%-24.4%]		5%	[3.9%-5.8%]
ES	**Never**	1297	5.3%	27.3%	25%	[23.0%-27.7%]	5.6%	13%	[11.6%-15.3%]
	0-2	12139	49.9%		30%	[29.4%-31.1%]		6%	[5.6%-6.4%]
	3-4	3957	18.7%		23%	[21.7%-24.3%]		4%	[3.3%-4.6%]
	5-6	5486	26.2%		17%	[16.2%-18.2%]		3%	[2.9%-3.8%]
FI	**0-2**	1944	26.2%	30.8%	35%	[33.1%-37.4%]	7.5%	11%	[9.6%-12.4%]
	3-4	3010	39.7%		32%	[30.3%-33.6%]		8%	[7.0%-8.9%]
	5-6	2697	34.0%		25%	[23.1%-26.3%]		4%	[3.6%-5.2%]
FR	**Never**	17	0.1%	24.6%	60%	[36.5%-83.1%]	8.6%	35%	[12.2%-57.5%]
	0-2	4977	36.8%		31%	[29.8%-32.4%]		12%	[10.9%-12.7%]
	3-4	6242	40.3%		23%	[21.8%-23.9%]		7%	[6.7%-7.9%]
	5-6	3706	22.8%		17%	[15.7%-18.1%]		5%	[4.0%-5.4%]
GR	**Never**	311	1.9%	18.8%	62%	[56.4%-67.2%]	7.1%	35%	[29.4%-39.9%]
	0-2	5410	44.2%		23%	[21.5%-23.7%]		9%	[8.0%-9.5%]
	3-4	3407	32.3%		16%	[14.7%-17.2%]		6%	[5.0%-6.6%]
	5-6	2248	21.6%		11%	[9.3%-11.9%]		3%	[2.1%-3.5%]

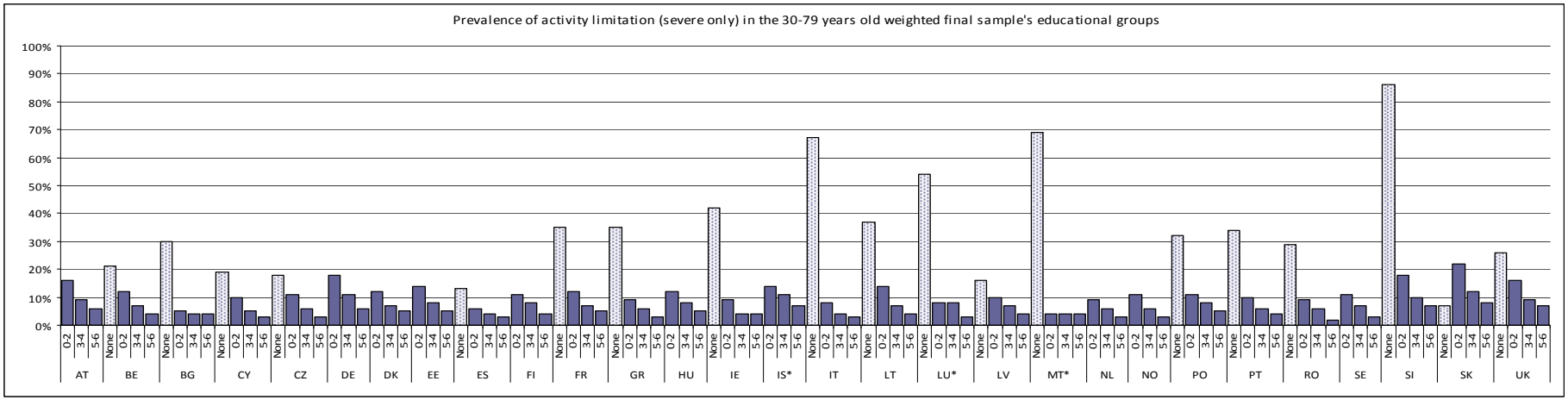
HU	**0-2**	3858	25.1%	33.1%	44%	[42.2%-45.3%]	9.2%	12%	[11.4%-13.5%]
	3-4	8562	56.7%		32%	[30.9%-32.8%]		8%	[7.7%-8.8%]
	5-6	2776	18.2%		22%	[20.1%-23.1%]		5%	[4.6%-6.3%]
IE	**Never**	67	0.7%	23.0%	62%	[50.6%-73.8%]	6.4%	42%	[30.3%-53.9%]
	0-2	3253	42.9%		29%	[27.5%-30.7%]		9%	[7.8%-9.7%]
	3-4	1900	28.0%		20%	[18.0%-21.6%]		4%	[3.1%-4.9%]
	5-6	2050	28.4%		13%	[11.6%-14.6%]		4%	[2.9%-4.6%]
IS*	***0-2***	*608*	*27.8%*	*17.7%*	*23%*	*[19.4%-26.1%]*	*10.6%*	*14%*	*[11.1%-16.6%]*
	3-4	*914*	*42.6%*		*19%*	*[16.1%-21.1%]*		*11%*	*[9.2%-13.3%]*
	5-6	*615*	*29.6%*		*12%*	*[9.4%-14.6%]*		*7%*	*[4.7%-8.7%]*
IT	**Never**	217	0.7%	25.9%	82%	[77.2%-87.3%]	6.6%	67%	[61.1%-73.6%]
	0-2	16724	53.7%		30%	[29.1%-30.4%]		8%	[7.4%-8.2%]
	3-4	11723	33.4%		21%	[20.1%-21.6%]		4%	[4.0%-4.8%]
	5-6	3831	12.2%		15%	[14.0%-16.2%]		3%	[2.9%-4.1%]
LT	**Never**	34	0.4%	26.2%	51%	[34.5%-68.1%]	7.7%	37%	[20.6%-53.0%]
	0-2	1881	20.9%		36%	[34.2%-38.5%]		14%	[12.0%-15.1%]
	3-4	4468	54.5%		27%	[25.7%-28.3%]		7%	[6.2%-7.6%]
	5-6	1820	24.2%		17%	[15.2%-18.6%]		4%	[2.8%-4.5%]
LU*	***Never***	*50*	*0.8%*	*23.5%*	*80%*	*[69.4%-91.4%]*	*7.2%*	*54%*	*[40.4%-68.0%]*
	0-2	*2584*	*39.8%*		*29%*	*[26.9%-30.4%]*		*8%*	*[7.2%-9.3%]*
	3-4	*2146*	*35.2%*		*21%*	*[19.4%-22.8%]*		*8%*	*[6.5%-8.7%]*
	5-6	*1711*	*24.1%*		*16%*	*[14.7%-18.2%]*		*3%*	*[2.5%-4.2%]*
LV*	**Never**	26	0.2%	35.4%	27%	[9.8%-43.9%]	6.9%	16%	[1.8%-29.8%]
	0-2	2021	19.7%		46%	[43.5%-47.8%]		10%	[8.5%-11.0%]
	3-4	4818	56.1%		36%	[34.2%-36.9%]		7%	[6.1%-7.6%]
	5-6	1949	23.9%		27%	[24.6%-28.5%]		4%	[3.1%-4.8%]
MT	***Never***	*99*	*1.4%*	*14.3%*	*89%*	*[82.7%-95.1%]*	*4.2%*	*69%*	*[59.4%-77.7%]*
	0-2	*4621*	*71.2%*		*15%*	*[14.0%-16.0%]*		*4%*	*[3.6%-4.8%]*
	3-4	*922*	*16.0%*		*12%*	*[9.8%-13.9%]*		*4%*	*[2.4%-4.8%]*
	5-6	*638*	*11.4%*		*9%*	*[7.1%-11.7%]*		*4%*	*[2.1%-5.0%]*
NL	**0-2**	2162	30.3%	29.1%	38%	[36.4%-40.5%]	5.8%	9%	[7.7%-10.1%]
	3-4	3186	39.6%		28%	[26.5%-29.7%]		6%	[5.0%-6.6%]
	5-6	2703	30.0%		22%	[20.0%-23.2%]		3%	[2.5%-3.8%]
NO	**0-2**	702	22.3%	18.2%	28%	[25.1%-31.8%]	6.0%	11%	[8.4%-12.9%]
	3-4	1845	46.4%		18%	[16.4%-19.9%]		6%	[4.7%-6.9%]
	5-6	1474	31.3%		10%	[8.4%-11.4%]		3%	[1.9%-3.5%]
PO	**Never**	27	0.1%	26.7%	62%	[43.7%-80.4%]	8.0%	32%	[14.0%-49.1%]
	0-2	4800	20.0%		34%	[32.8%-35.5%]		11%	[10.5%-12.3%]
	3-4	13715	63.7%		26%	[25.2%-26.7%]		8%	[7.3%-8.2%]
	5-6	3047	16.2%		19%	[17.7%-20.5%]		5%	[4.6%-6.2%]
PT	**Never**	1140	11.1%	34.9%	52%	[49.5%-55.3%]	10.8%	34%	[31.7%-37.2%]
	0-2	5736	67.7%		35%	[34.2%-36.7%]		10%	[9.5%-11.1%]
	3-4	775	10.7%		24%	[20.6%-26.6%]		6%	[4.0%-7.3%]
	5-6	755	10.5%		19%	[15.8%-21.4%]		4%	[2.8%-5.7%]
RO	**Never**	206	1.5%	25.1%	47%	[40.0%-53.6%]	7.5%	29%	[23.0%-35.4%]
	0-2	4946	35.4%		29%	[27.6%-30.1%]		9%	[8.2%-9.8%]
	3-4	6282	52.6%		23%	[22.4%-24.5%]		6%	[5.8%-7.0%]
	5-6	1193	10.4%		15%	[13.2%-17.3%]		2%	[1.3%-2.9%]
SE	**0-2**	1017	17.8%	16.2%	22%	[19.0%-24.0%]	6.4%	11%	[9.3%-13.2%]
	3-4	2857	50.4%		17%	[16.0%-18.8%]		7%	[5.9%-7.7%]
	5-6	1787	31.8%		12%	[10.5%-13.5%]		3%	[2.1%-3.7%]
SI	**Never**	36	0.4%	28.4%	94%	[87.0%-101%]	11.6%	86%	[74.6%-97.3%]
	0-2	1707	23.6%		38%	[35.5%-40.1%]		18%	[16.0%-19.6%]
	3-4	3947	56.3%		27%	[25.9%-28.7%]		10%	[9.3%-11.1%]
	5-6	1415	19.7%		18%	[16.4%-20.4%]		7%	[5.9%-8.6%]
SK	**Never**	23	0.2%	41.6%	21%	[4.5%-38.0%]	12.8%	7%	[-3.3%-18.3%]
	0-2	975	10.1%		54%	[51.1%-57.3%]		22%	[19.8%-25.1%]
	3-4	6773	72.6%		42%	[40.7%-43.0%]		12%	[11.5%-13.1%]
	5-6	1520	17.0%		34%	[31.2%-35.9%]		8%	[6.9%-9.7%]
UK	**Never**	11	0.1%	22.8%	55%	[25.1%-84.0%]	10.0%	26%	[0.2%-52.2%]
	0-2	2825	24.5%		31%	[29.7%-33.2%]		16%	[14.6%-17.3%]
	3-4	4967	46.0%		22%	[21.3%-23.6%]		9%	[8.6%-10.2%]
	5-6	3145	29.4%		17%	[15.2%-17.8%]		7%	[5.7%-7.4%]

* Countries whose samples display a distribution by education level that deviates from distribution in the general population.

Figure 5: Percentages of education levels among the 30-79s and prevalence of activity limitations (all levels of severity combined and severe levels only) in EU-SILC weighted final sample







*Countries whose samples display a distribution by education level that deviates from distribution in the general population.

Summary

The EU-SILC survey exhibits disparities in survey methods between countries, such as the choice of interview procedure and the use of additional data sources. The distinctive feature of EU-SILC is that the data are harmonized after collection, unlike in other European surveys such as SHARE (Survey of Health, Ageing and Retirement in Europe) and ESS (European Social Survey), for which the data are harmonized before collection.

In consequence, there are wide variations in non-response rates and in the collection of certain variables. The end result is a reduced sample for health variables, particularly in countries that rely on register data, conduct telephone interviews, and prohibit the collection of the health variable by proxy. We also observe missing data for education and occupation variables, particularly for the oldest ages. A study comparing social inequalities in health by education level should be confined to respondents who have supplied information on all these variables, a constraint that further reduces sample size. In Denmark, Finland, Iceland, the Netherlands, Norway, Sweden, and Slovenia, more than one-third of the EU-SILC samples (based on registers) cannot be used to study activity limitations because of non-response from persons selected at random (non-response or response by a proxy who is not allowed to answer questions on health variables for someone else).

However, sample attrition does not invariably undermine the quality of the observations collected. That depends on the specific attrition factors. If they concern health and/or social status, they can alter the scale of prevalence of activity limitations in the groups surveyed. They can minimize it if non-respondents are more "limited" (refusal/cannot take part in survey) or overestimate it if non-respondents are—on the contrary—less limited (young economically active people who are hard to reach). Two effects operate simultaneously: (1) the selection effect, if non-respondents and non-participants differ significantly from respondents in terms of health or education; (2) the distortion effect, which can alter prevalence in the total population if the sample structure differs significantly from that of the population studied, by shifting the mean value towards the value of the best-represented category.

The selection effect due to non-participants in EU-SILC is not measurable, as their social status and health are not known. The selection effect due to non-respondents in the final sample can be assessed, as some of the information—particularly on social status—can be gathered from the "household" database even if the person has not taken part in the individual survey. Distortion can be measured by comparing the social structures of the final sample with that of the complete EU-SILC database and the reference data from Eurostat being based on the Labor Force Survey. Meanwhile, the latter is also subject to varying response rates so are not real gold standard to assess the quality of the distribution.

While thirteen countries display a sample structure resembling that of their total population, others diverge from it to a fairly significant extent. In some countries, low education levels are rather substantially under-represented (SE, IS, MT, UK) or less strongly so (CY, ES, FR, GR, NL, SK). The Belgian sample tends to over-represent the highest education level, but the share of the lowest level is the same as in the total population. In Luxembourg, the lowest education level is largely over-represented and—as in Iceland and Malta—the sample distortions lead to one of the education levels being under- or over-represented by more than ten percentage points.

As regards the sample's age structure, the youngest age groups tend to be under-represented. The determinants of non-response to the individual survey are the fact of being relatively young and economically active (and so impossible to reach), a probable effect of the many individual surveys conducted by telephone. However, we find that the response rate for the education, occupation, and health variables can be rather low among the over-80s. This is partly due to the fact that, for the very elderly, the answers are more often given by a proxy who may have little knowledge of the

person's education level and who, moreover, is banned in many countries from answering the health question on the person's behalf.

On balance, we find that the countries relying on registers and telephone surveys display very low participation rates in the individual survey and very low response rates for the education and health variables. Yet the massive attrition of the respondent population does not significantly alter the structures by age, education or socio-occupational category. We note an under-representation of the youngest persons (not reached) and the very elderly (when a proxy needs to answer). The result presumably reflects an under-estimation of activity limitations among the over-79s, as proxy responses are very likely due, at these ages, to the difficulty of answering in person. At the youngest ages, the effect of a proxy response or non-participation probably most often reflects lack of time or a refusal to participate, without indicating potential disabilities. In all likelihood, therefore, the effect is related to age but also to social status. The groups under-represented at very old ages will probably cause an under-estimation of the prevalence of activity limitations, and the effect in terms of prevalence in the general population will be all the stronger as the under-representation will involve the social groups with the highest proportions of activity limitations, i.e., the low-education groups.

Given their structure, it is preferable to exclude from the study the countries that largely under- or over-represent certain groups. This applies to Malta, Luxembourg, and Iceland. Moreover, to minimize uncertainty, we recommend confining the study to ages 30-79, for which the fullest information regarding education is available. More generally, we recommend singling out the countries for which one may suspect a bias given the distortion of the samples' social structures.

A deeper analysis of the prevalences of activity limitations and their differentials will be the subject of a new report. On the basis of the initial figures supplied here, we observe that the expected social gradient is consistently visible, as the prevalences of activity limitations decrease when the education level increases. Fuller analyses will seek to identify trends in these differentials at European level. We shall attempt to quantify an education/health link in the countries and to examine whether it matches the findings of other studies, particularly the studies that emphasize a specific effect of health and social-protection policies implemented in different regions of the European Union.

Appendix

Table 12: Summary table of EU-SILC data quality (the asterisk * denotes a problem: low response rates, sample attrition, sample representativeness, distortion, impact of distortion on prevalence)

	Response rate at individual level	Loss in final sample	Missing data			Representativeness			Distortions			Adjustment of prevalences by distribution	
			Education	Occupation	Health	Education	Occupation	Age	Education	Occupation	Age	Limitation	Severe limitation
AT	*												
BE	*			*			*						
BG	*												
CY							*						
CZ		*						*	*	*	*		
DE	*						*						
DK	*	*	80 +*	*					*	*	*		
EE	*				*								
ES						*							
FI		*							*	*	*		
FR				*, 80 +*									
GR						*							
HU													
IE						*		*				*	
IS	*	*				*			*	*	*		
IT													
LT					*								
LU*	*					*	*					*	
LV	*												
MT	*			80 +*		*							
NL		*		*					*	*	*		
NO	*	*		*			*		*	*	*		
PL	*		*								*		
PT													
RO													
SE	*	*	80 +*	*		*			*	*	*		
SI	*	*							*	*	*		
SK							*	*					
UK	*		*			*							

Table 13: ISCED 1997 education levels

Level	English	French	Main characteristics[7]
0	Pre-primary level	Education préprimaire	*Initial stage of organized instruction are designed primarily to introduce very young children to a school-type environment, i.e., to provide a bridge between the home and a school-based atmosphere.*
1	Primary level	Enseignement primaire	*Programmes at level 1 are normally designed on a unit or project basis to give students a sound basic education in reading, writing and mathematics along with an elementary understanding of other subjects such as history, geography, natural science, social science, art and music. In some cases religious instruction is featured.*
2	Lower secondary level	Premier cycle de l'enseignement secondaire	*Typically designed to complete the provision of basic education which began at ISCED level 1. In many, if not most countries, the educational aim is to lay the foundation for lifelong learning and human development on which countries may expand, systematically, further educational opportunities. The programmes at this level are usually on a more subject-oriented pattern using more specialized teachers (as distinct from level 1).*
3	Upper secondary level	Enseignement secondaire	*This level of education typically begins at the end of full-time compulsory education for those countries that have a system of compulsory education. More specialization may be observed at this level than at ISCED level 2. Minimum entrance requirements are the completion of level 2 or demonstrable ability to handle programmes at this level.*
4	Post-secondary non-tertiary level	Enseignement postsecondaire non supérieur	*Straddles the boundary between upper- secondary and post-secondary education from an international point of view. Designed to prepare students for studies at level 5 who, although having completed ISCED level 3, did not follow a curriculum which would allow entry to level 5, i.e., pre-degree foundation courses or short vocational programmes.*
5	First stage of tertiary level	Premier cycle de l'enseignement supérieur	*Educational content more advanced than those offered at levels 3 and 4. Programmes must have a cumulative theoretical duration of at least 2 years. Does not lead directly to the award of an advanced research qualification (level 6).*
6	Second stage of tertiary level	Second cycle de l'enseignement supérieur	*Reserved for tertiary programmes which lead to the award of an advanced research qualification. Programmes are therefore devoted to advanced study and original research and are not based on course-work only.*

[7] UNESCO, International Standard Classification of Education (ISCED 1997), http://www.uis.unesco.org/Library/Documents/isced97-en.pdf (French version: http://www.uis.unesco.org/Library/Documents/isced97-fr.pdf) (retrieved 22/10/2012).

Table 14: CITP-88 - Sub-major groups

	ENGLISH	FRENCH
0	**Armed forces occupations**	**Professions militaires**
01	Commissioned armed forces officers	Officiers des forces armées
02	Non-commissioned armed forces officers	Sous-officiers des forces armées
03	Armed forces occupations, other ranks	Autres membres des forces armées
1	**Managers**	**Managers**
11	Chief executives, senior officials and legislators	Directeurs généraux, cadres supérieurs et membres de l'Exécutif et des corps législatifs
12	Administrative and commercial managers	Managers de services administratifs et commerciaux
13	Production and specialized services managers	Managers, production et services spécialisés
14	Hospitality, retail and other services managers	Managers de l'hôtellerie, la restauration, le commerce de détail et de gros et autres services
2	**Professionals**	**Professions intellectuelles, scientifiques et artistiques**
21	Science and engineering professionals	Spécialistes des sciences techniques
22	Health professionals	Spécialistes de la santé
23	Teaching professionals	Spécialistes de l'enseignement
24	Business and administration professionals	Spécialistes en gestion et administration d'entreprises
25	Information and communications technology professionals	Spécialistes des technologies de l'information et des communications
26	Legal, social and cultural professionals	Spécialistes de la justice, des sciences sociales et de la culture
3	**Technicians and associate professionals**	**Professions intermédiaires**
31	Science and engineering associate professionals	Professions intermédiaires des sciences et techniques
32	Health associate professionals	Professions intermédiaires de la santé
33	Business and administration associate professionals	Professions intermédiaires, finance et administration
34	Legal, social, cultural and related associate professionals	Professions intermédiaires des services juridiques, des services sociaux et assimilés
35	Information and communications technicians	Techniciens de l'information et des communications
4	**Clerical support workers**	**Employés de type administratif**
41	General and keyboard clerks	Employés de bureau
42	Customer services clerks	Employés de réception, guichetiers et assimilés
43	Numerical and material recording clerks	Employés des services comptables, financiers, de paie et assimilés et magasiniers
44	Other clerical support workers	Autres employés de type administratif
5	**Service and sales workers**	**Personnel des services directs aux particuliers, commerçants et vendeurs**
51	Personal service workers	Personnel des services directs aux particuliers
52	Sales workers	Commerçants et vendeurs
53	Personal care workers	Personnel soignant
54	Protective services workers	Personnel des services de protection et de sécurité
6	**Skilled agricultural, forestry and fishery workers**	**Agriculteurs et ouvriers qualifiés de l'agriculture, de la sylviculture et de la pêche**
61	Market-oriented skilled agricultural workers	Agriculteurs et ouvriers qualifiés de l'agriculture commerciale
62	Market-oriented skilled forestry, fishery and hunting workers	Professions commerciales qualifiées de la sylviculture, de la pêche et de la chasse
63	Subsistence farmers, fishers, hunters and gatherers	Agriculteurs, pêcheurs, chasseurs et cueilleurs de subsistance
7	**Craft and related trades workers**	**Métiers qualifiés de l'industrie et de l'artisanat**
71	Building and related trades workers, excluding electricians	Métiers qualifiés du bâtiment et assimilés, sauf électriciens
72	Metal, machinery and related trades workers	Métiers qualifiés de la métallurgie, de la construction mécanique et assimilés
73	Handicraft and printing workers	Métiers qualifiés de l'artisanat et de l'imprimerie
74	Electrical and electronic trades workers	Métiers de l'électricité et de l'électrotechnique
75	Food processing, wood working, garment and other craft and related trades workers	Métiers de l'alimentation, du travail sur bois, de l'habillement (y compris l'ameublement) et autres métiers qualifiés de l'industrie et de l'artisanat

8	**Plant and machine operators, and assemblers**	**Conducteurs d'installations et de machines, et ouvriers de l'assemblage**
81	Stationary plant and machine operators	Conducteurs de machines et d'installations fixes
82	Assemblers	Ouvriers de l'assemblage
83	Drivers and mobile plant operators	Conducteurs de véhicules et d'engins lourds de levage et de manœuvre
9	**Elementary occupations**	**Professions élémentaires**
91	Cleaners and helpers	Aides de ménage
92	Agricultural, forestry and fishery labourers	Manœuvres de l'agriculture, de la pêche et de la sylviculture
93	Labourers in mining, construction, manufacturing and transport	Manœuvres des mines, du bâtiment et du génie civil, des industries manufacturières et des transports
94	Food preparation assistants	Collaborateurs en restauration rapide
95	Street and related sales and service workers	Vendeurs ambulants et autres travailleurs des petits métiers des rues et assimilés
96	Refuse workers and other elementary workers	Eboueurs et autres travailleurs non qualifiés

Co-funded by 10 Member States, the European Commission, DG SANCO and two French institutions: DREES and CNSA.

Contact EHLEIS:
Jean Marie ROBINE, INSERM
Université Montpellier II / U710 - MMDN
Place Eugène Bataillon, bat 24 - CC105
34095 Montpellier Cedex 05, France.

Tel: +33 (0) 467 14 33 85
Fax: +33 (0) 467 14 92 95
Email: jean-marie.robine@inserm.fr

www.eurohex.eu