A note on the replication of the European Socio-economic Classification (ESeC) in the EU Statistics on Income and Living Conditions (EU-SILC)

Tim Goedemé

December 2019

INET Oxford Working Paper No. 2019-17

Employment, Equity & Growth

A note on the replication of the European Socio-economic Classification (ESeC) in the EU Statistics on Income and Living Conditions (EU-SILC)

Tim Goedemé, PhD

Institute for New Economic Thinking at the Oxford Martin School

Department of Social Policy and Intervention – University of Oxford

Associate Member of Nuffield College

Herman Deleeck Centre for Social Policy – University of Antwerp

tim.goedeme@spi.ox.ac.uk

Abstract

The European Socio-Economic Classification (ESeC) is quite popular in social science research, especially as an indicator of social class. The EU Statistics on Income and Living Conditions (EU-SILC) provides the opportunity to study the relation between ESeC and a long list of socio-economic variables (including income, employment, material deprivation, …) in a comparative and longitudinal perspective. However, to the best of my knowledge, very few people work with ESeC in EU-SILC. Therefore, in this research note, I document the data quality of EU-SILC in terms of measuring ESeC, and propose a slightly modified version of computing ESeC compared to the routines one can find on the GESIS webiste. More in particular, I look at the availability and comparability of the variables that are required to construct ESeC, document some issues with non-response bias and the comparability of the selected respondent countries, and provide list of general recommendations when using ESeC in EU-SILC, as well as a detailed list of issues by country. An excel file with background tables and a STATA do-file to construct ESeC for all EU-SILC countries in EU-SILC 2014-2017 accompany this paper.

1   Introduction[1]

In this note, I discuss some data quality issues in the construction of the European Socio-economic Classification (ESeC) in the most important survey for comparative research on income and living conditions in the European Union, EU-SILC. The European Socio-Economic Classification is a categorical social class schema that was developed more than a decade ago to facilitate comparative research on social class in Europe (Rose and Harrison, 2007). The class schema is based on the type of employment relations in which one is involved and occupation. More specifically, it makes a distinction between the

---

self-employed and employees, and within the self-employed between those with many and those with few or no employees. Further, at least in its implementation, employees with a supervisory role are distinguished from those without such a role. Those voluntary out of paid employment are considered a separate category. On top of this, occupation functions as a discriminator, assigiaring people to one of nine classes, depending on the employment relation they have. Occupation is measured on the basis of the International Standard Classiification of Occupations (ISCO)[2]. To the best of my knowledge, there are not many papers that use EU-SILC for social class analysis, although (1) EU-SILC contains many variables that are relevant to researchers of social class (e.g. on labour market position and outcomes, earnings, household incomes, material deprivation and housing conditions); and (2) for many analyses of poverty, inequality, labour market conditions and living standards, social class functions probably as an important determinant. Nonetheless, EU-SILC contains the key variables to reconstruct an ESeC social class schema, even though with some noticeable caveats. With this research note, I aim to provide some more insight into the suitability of the EU-SILC data for social class research, from a data quality perspective.

To operationalise ESeC, I start from the STATA do-files that are made available online by the German GESIS institute[3]. These do-files try to make the best of the EU-SILC variables to reconstruct ESeC, based on the instructions by Rose and Harrison (2007). I update the do-files to reflect the situation of the most recent EU-SILC years (up to EU-SILC 2017), and evaluate the implementation of EseC for all available countries in the currently available EU-SILC User Database, in the period from 2004 till 2017. We restrict our focus to an individual (rather than a household) social class measure, and to the population in paid employment at working age. To the best of my knowledge, there are no similar evaluations available. We find that users of EU-SILC should be careful when 'quickly' comparing results by social class across countries or across time, and take changes and cross-national variations in data availability and quality seriously, or they might end up taking survey errors for 'substantive findings'. This note is part of a broader project of the Employment Equity and Growth programme at the Institute for New Economic Thinking at the Oxford Martin School, funded by Citi, in which we assess the relation between social class and income inequality in rich countries. The STATA do-files and detailed outputs in excell are available as an online appendix to this note[4].

In what follows, I start with a very brief introduction to ESeC and how it is reconstructed in EU-SILC. Subsequently, I discuss overall data availability and 'official' breaks in series. In the following section, I look consecutively at the key variables that define social class, as available in EU-SILC. Thereafter, I assess overall data availability of ESeC in EU-SILC and briefly describe some data problems in terms of non-random missing cases and fluctuating shares of the self-employed in the middle and upper classes. Next, I present the results of a slightly modified version of the GESIS code to reconstruct ESeC in EU-SILC, which I believe improves the validity of the scale. I conclude with a brief summary of the main overall issues and a list of remarks by country to keep in mind when studying social class with EU-SILC.

2   The European Socio-Economic Classification in EU-SILC

The European Socio-Eonomic Classification builds on multiple layers of classifying someone's professional position in the economy to subdivide people into 10 social classes. These layers include the employment

---

[2] See https://www.ilo.org/public/english/bureau/stat/isco/ (last consulted 19 June 2019).

[3] The do-files were developed by Anika Herter and Heike Wirth and can be downloaded from https://www.gesis.org/en/gml/european-microdata/eu-silc/ (last consulted 19 June 2019)

[4] https://timgoedeme.com/tools/esec-in-eu-silc/.

position, the size of the firm (for employers), the occupation, and, for employees, the supervisory role. Each layer co-determines in which class someone ends up, but classification does not follow a classical 'tree structure': people who are in a different labour market position in some ways (e.g. self-employed vs. employees), may still end up in the same social class, depending on their occupation. This is further illustrated in the figure below.

In EU-SILC, the required information is not available in exactly the same way as in other surveys, such as the European Social Survey (ESS), but there are variables that allow for a reasonable approximation. The second figure below illustrates how the GESIS code makes use of the available EU-SILC variables to reconstruct ESeC on the basis of ISCO-88. With ISCO-08 (for most countries introduced in EU-SILC from 2011 onwards) the flowchart should be somewhat modified, in the sense that some large employers are classified in classes three and five and some self-employed are classified in class three. In addition, with ISCO-08 some non-supervisory employees are classified among the small farmers, and no occupation in this group is any longer classified as a higher grade blue collar worker. In this note, as well as in the schema below, I do not consider the class consisting of those who are not in paid employment. For convenience, the nine social classes are sometimes collapsed into a more limited number of classes. Following Rose and Harrisson (Z007), also in this paper I sometimes collapse them into three: the upper class or salariat (ESeC classes one and two); the middle, or intermediate, class (ESeC classes 3-6) and the lower class or working class (classes 7-9).

Figure 1. The conceptual derivation of the European Socio-Economic Classification
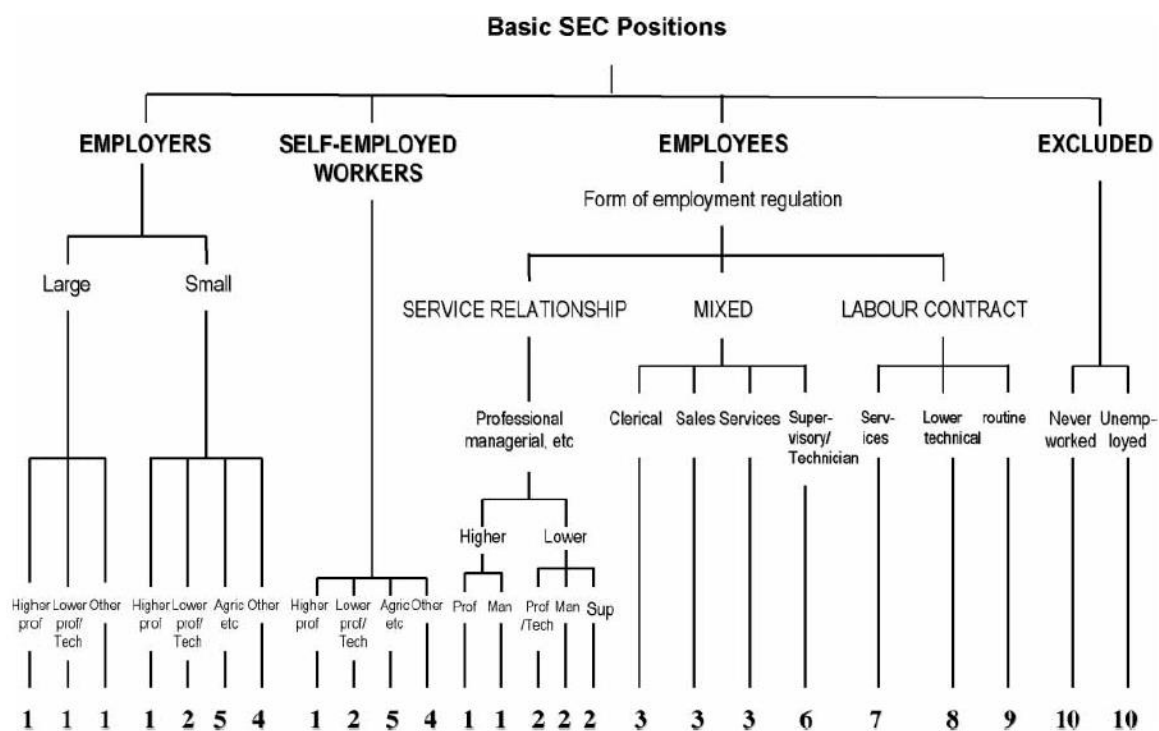
Source: Rose and Harrison (2007, p. 471)

Figure 2. Schematic representation of the GESIS code to reconstruct ESeC in EU-SILC

Source: own compilation on the basis of GESIS STATA do-files.

3   Availability of variables and breaks in series

3.1   Overall availability of data and official breaks in series

At the time of writing, we had EU-SILC data available for almost all 32 EU-SILC countries from 2004 to 2017 (interview years). Note all of these are available for the entire period, which applies especially to the newer Member States.

Table 1. Availability of EU-SILC data (latest update of April 2019), 1 represents 'yes'

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
AT	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BE	1	1	1	1	1	1	1	1	1	1	1	1	1	1
BG	0	0	0	1	1	1	1	1	1	1	1	1	1	1
CH	0	0	0	1	1	1	1	1	1	1	1	1	1	0
CY	0	1	1	1	1	1	1	1	1	1	1	1	1	1

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
CZ	0	1	1	1	1	1	1	1	1	1	1	1	1	1
DE	0	1	1	1	1	1	1	1	1	1	1	1	1	1
DK	1	1	1	1	1	1	1	1	1	1	1	1	1	1
EE	1	1	1	1	1	1	1	1	1	1	1	1	1	1
EL	1	1	1	1	1	1	1	1	1	1	1	1	1	1
ES	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FI	1	1	1	1	1	1	1	1	1	1	1	1	1	1
FR	1	1	1	1	1	1	1	1	1	1	1	1	1	1
HR	0	0	0	0	0	0	1	1	1	1	1	1	1	1
HU	0	1	1	1	1	1	1	1	1	1	1	1	1	1
IE	1	1	1	1	1	1	1	1	1	1	1	1	1	1
IS	1	1	1	1	1	1	1	1	1	1	1	1	1	0
IT	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LT	0	1	1	1	1	1	1	1	1	1	1	1	1	1
LU	1	1	1	1	1	1	1	1	1	1	1	1	1	1
LV	0	1	1	1	1	1	1	1	1	1	1	1	1	1
MT	0	0	0	1	1	1	1	1	1	1	1	1	1	1
NL	0	1	1	1	1	1	1	1	1	1	1	1	1	1
NO	1	1	1	1	1	1	1	1	1	1	1	1	1	1
PL	0	1	1	1	1	1	1	1	1	1	1	1	1	1
PT	1	1	1	1	1	1	1	1	1	1	1	1	1	1
RO	0	0	0	1	1	1	1	1	1	1	1	1	1	1
RS	0	0	0	0	0	0	0	0	0	1	1	1	1	1
SE	1	1	1	1	1	1	1	1	1	1	1	1	1	1
SI	0	1	1	1	1	1	1	1	1	1	1	1	1	1
SK	0	1	1	1	1	1	1	1	1	1	1	1	1	1
UK	0	1	1	1	1	1	1	1	1	1	1	1	1	1

Note: Cells that represent 'official' breaks in series are shaded in dark grey.

For just a few countries, Eurostat announces on its website 'official' breaks in series. The listed breaks in series for the at-risk-of-poverty rate are[5]: Sweden in 2008; Bulgaria, Luxembourg and the Netherlands in 2016; and United Kingdom in 2017. It is somwhat surprising that other changes in data collection or weighting procedures are not counted as breaks in series (e.g. the change in underlying data source for the UK in 2012[6], the change in weighting schemes in Belgium since 2012; the increased used of register data for collecting income information for a range of countries; …).

---

[5] At-risk-of-poverty rate by poverty threshold, age and sex - EU-SILC survey [ilc_li02], last updated 4 April 2019 and extracted 23 April 2019 from the Eurostat online database.
[6] In 2012 the Family Resources Survey replaced the General Lifestyle Survey as the main source for EU-SILC, see for instance, https://beta.ukdataservice.ac.uk/datacatalogue/series/series?id=200015#!/faqs (last accessed 19 June 2019).

3.2   Breaks in individual variables required for ESeC

To construct an ESeC variable in EU-SILC, the GESIS code makes use of the following variables:

- Variable PL050 / PL051 which contains the ISCO-88, respectively the ISCO-08 classification.
- Variable PL040 which indicates the labour market status in four categories, which allows to distinguish the self-employed from the employees.
- Variable PL130, which describes the size of the local economic unit.
- Variable PL150, which describes whether some supervises others at work or not.

The overall fitness of each of these variables in terms of their validity and measurement error is outside the scope of this paper. In what follows, I focus primarily on consistency over time and across countries and point to several issues that (1) act as a warning for using the data for particular countries or certain years; (2) clearly undermine the comparability across countries.

3.2.1   Occupation: ISCO-88 (PL050) vs. ISCO-08 (PL051)

An important factor to determine someone's social class in ESeC is the occupational group to which one belongs. EU-SILC provides, in principle, occupational class at the two-digit level, but for some countries the categories are available only at a more aggregated level. There are several limitations, though: (1) in principle ESeC requires a breakdown at the three-digit level; (2) in 2011 (in most countries) ISCO-08 replaced ISCO-88 as the occupation variable. This is in line with the global revision of the International Standard Classification of Occupations. An important change is that ISCO-08 is more detailed at the two-digit level than ISCO-88 was, but the change causes a break in series, as we will illustrate below. Fortunately, for most countries, both ISCO-88 and ISCO-08 are available for EU-SILC 2011, which helps to gain more insight into the effect of the change in occupational classification. The exceptions are BG, FI and IE (ISCO-88 missing in 2011). In contrast, in the case of CH, EE and ES both classifications are available for 2010 and 2011.

The level of detail with which the data are provided varies across countries, and in some countries additional changes have taken place. ISCO-88 typically is available with 27 categories, while only 26 categories are present in Belgium (except 2006 and 2008), Germany (except 2005), France, Iceland and the United Kingdom as well as in Slovakia for some years. Only 25 categories are present in Portugal. ISCO-08 is typically available in 40 to 43 categories, with somewhat more variation in the exact number across time and countries, probably as a result of a varying prevalence of these categories across countries and, in some cases, mistakes. For instance, the varying number of categories for Belgium is for several years clearly due to categories with a very low prevalence (less than 10 observations on average in the subcategories of the 'Skilled Agricultural, Forestry and Fishery Workers'), but for 2011 some categories are not present even though they have a reasonable prevalence in other years. Also for other countries, it seems to be primarily the categories with a very low number of observations that tend to drop out in some years. In contrast, for Germany the number is persistently

Quite a few countries show a more strongly deviating number of categories, though. In the case of Malta, ISCO-88 and ISOC-08 are only available in 10 categories (except for 2007). Furthermore, in Germany only 38 categories are used for ISCO-08, and this has dropped to nine categories since 2015. Similarly, in Slovenia only 10 categories are in use since 2014. In addition, in Ireland and in Slovakia (until 2014) considerably fewer categories are used, varying between 25 and 27. A comparison of the two variables for the 2011 Slovak data, shows a considerable mismatch between PL050 and PL051, implying that the

lower number of ISCO-08 categories is not due to a continued use of ISCO-88 rather than ISCO-08. For Ireland, a similar analysis is unfortunately not possible. Given that in the case of Slovakia all 40+ categories are used from 2015 onwards, it most likely collapsed some categories together, but it is unclear how this was exactly done. For Iceland, occupational category is not available since SILC 2014. Several countries have notified Eurostat of some smaller problems for particular years: In the Belgian 2005 SILC armed forces (code 01) have been coded by mistake as 11; for Hungary there were (non-specified) problems with variable PL050 in SILC 2005. Further, in the Netherlands a relatively small number of observations (378) is in the unique position of having received code 4, 7 or 8 in 2013-2017, which is not used in any other country/year.

Table 2. Number of categories available in PL051 (ISCO-08) in countries with inconsistencies

	2010	2011	2012	2013	2014	2015	2016	2017
DE	0	38	38	38	38	9	9	9
IE	0	25	26	25	25	25	25	25
IS	38	39	40	39	0	0	0	0
MT	0	10	10	10	10	10	10	10
NL	0	41	41	44	44	44	45	44
SI	0	41	41	41	10	10	10	10
SK	0	27	27	27	27	41	43	43

Apart from these limitations, there are considerable variations in the share of the sample for which ISCO-88 / ISCO-08 is available. The conditions for collecting the variable have changed slightly over time, but in general, it has to be collected for everyone who has worked at some point. Therefore, the share of the sample for which the information is available probably varies in line with this requirement, varying between 80 and 90 per cent of the sample, on average. An important exception is Slovenia, with a much lower average percentage of respondents and Bulgaria, Finland, Norway and Romania for some years. More worryingly, there are quite a few countries where the percentage suddenly drops or increases in some years, as can be seen from the table below.

Table 3. Proportion of persons with a non-missing value for ISCO-88 (PL050) countries with at least one 5-percentage point drop/increase in the proportion respondents among the sample aged [18-65[

	2004	2005	2006	2007	2008	2009	2010	2011
BE	0.79	0.82	0.85	0.85	0.84	0.84	0.85	0.80
BG	.	.	.	0.56	0.87	0.88	0.87	0.00
CH	.	.	.	0.85	0.80	0.78	0.83	0.87
DK	0.82	0.83	0.84	0.84	0.79	0.88	0.88	0.88
ES	0.71	0.82	0.82	0.85	0.87	0.86	0.86	0.83
FI	0.68	0.76	0.79	0.79	0.81	0.95	0.92	0.00
IS	0.84	0.86	0.88	0.82	0.88	0.91	0.91	0.90
NL	.	0.76	0.81	0.82	0.81	0.88	0.87	0.86
NO	0.78	0.72	0.78	0.79	0.80	0.83	0.88	0.82
SI	.	0.63	0.64	0.64	0.66	0.77	0.78	0.78

Table 4. Proportion of persons with a non-missing value for ISCO-08 (PL051), countries with at least one 5-percentage point drop/increase in the proportion respondents among the sample aged [18-65[

	2011	2012	2013	2014	2015	2016	2017
DK	0.90	0.88	0.87	0.85	0.85	0.89	0.82
EE	0.83	0.84	0.85	0.87	0.88	0.89	0.89
HR	0.81	0.79	0.82	0.82	0.84	0.83	0.80
HU	0.86	0.86	0.86	0.86	0.86	0.87	0.82
LU	0.88	0.89	0.88	0.88	0.90	0.76	0.74
MT	0.84	0.85	0.85	0.85	0.87	0.87	0.91
NO	0.80	0.88	0.90	0.73	0.85	0.86	0.87
RO	0.83	0.82	0.81	0.62	0.81	0.81	0.80
SE	0.82	0.83	0.84	0.86	0.86	0.86	0.87

3.2.2 Employment status (PL040 and PL031)

Another important ingredient for constructing ESeC regards someone's labour market status, making a distinction between employees, the self-employed and others. The main variable to do so is Status in Employment (PL040), which has to be collected for all current household members aged 16 and over, either through an interview, proxy or register. It should be collected for those currently working, and on the last job held for those currently out of work. It takes four values:

1. Self-employed with employees
2. Self-employed without employees
3. Employee
4. Family worker (in line with GESIS practice, we treat family workers as employees, even though one can wonder whether their security level is more in line with self-employed without employees, rather than employees)

'Self-defined current economic status' (PL031) appears to have been collected for somewhat more persons, so can be used in combination with PL040 to identify employees, the self-employed and others. However, PL031 is available for most countries since 2009 only, so it cannot be used for the entire period of observation. Given the transition from ISCO-88 to ISCO-08 in 2011, it seems advisable to start using PL031 also from 2011 onwards, to avoid any other break in series in the pre-2011 period.

For some countries, the percentage of the sample at active age for which values are recorded for PL040 have changed in a strange way over time. A (more than) ten percentage point increase in the proportion of people at active age that are an employee or self-employed from one year to the next is extremely unlikely, but that is what we observe in the data for some country/years. These variations are also likely to affect the composition of social classes and may bias findings regarding, for instance, within and between social class inequalities. For Bulgaria and Ireland, this implies missing just the first two years, but for Germany, the Netherlands, Norway and Slovenia it implies that the data can only be used from 2009 or 2010 onwards. This means that for these countries, we can only look in a proper way at the period 2011-2017 (see the variable on ISCO coding below). Even though, as we will show below, for the latter three countries, which are so-called 'selected respondent countries', these figures are only relevant insofar they affect selected respondents.

Table 5. Percentage of the sample at working age ([18-65[years old) for which PL040 is not missing

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
BG	.	.	.	0.56	0.87	0.88	0.87	0.86	0.86	0.86	0.87	0.87	0.87	0.88
DE	.	0.90	0.92	0.77	0.77	0.78	0.93	0.93	0.94	0.93	0.93	0.92	0.93	0.93
FI	0.68	0.77	0.78	0.80	0.82	0.95	0.92	0.87	0.85	0.88	0.89	0.89	0.89	0.89
IE	0.82	0.73	0.89	0.89	0.89	0.88	0.88	0.87	0.87	0.87	0.85	0.87	0.86	0.86
NL	.	0.76	0.83	0.83	0.85	0.92	0.91	0.91	0.91	0.90	0.90	0.91	0.91	0.92
NO	0.81	0.80	0.80	0.80	0.81	0.84	0.90	0.92	0.91	0.91	0.75	0.87	0.89	0.89
SI	.	0.62	0.67	0.67	0.69	0.81	0.81	0.81	0.81	0.81	0.81	0.82	0.82	0.83

This number can be adjusted by including also those who report currently being an employee or self-employed in PL031: if someone states that s/he is currently an employee or self-employed we may assume that a missing value for PL040 is a mistake. However, filling up missing values based on PL031 has only very marginal effects on the size of the available sample: only in Spain for a few years, this increases the sample size with between one and two percentage points. In all other countries, it adds a very low number of observations, amounting to an increase of less than one percentage point, and mostly less than 0.3 percentage points.

3.2.3   Size of local unit (PL130)

In the case of self-employed, a distinction is made between small and larger 'local economic units'. This information is provided by variable PL130, with the following response categories:

- 1 – 10: Exact number (if between 1 and 10)
- 11: Between 11 and 19 persons
- 12: Between 20 and 49 persons
- 13: 50 persons and more
- 14: Do not know but less than 11 persons
- 15: Do not know but more than 10 persons

For the creation of ESeC, we do not need this level of detail, but just need the variable to make a distinction between those working at a unit with fewer than 10 employees, and units with at least 10 employees. In the case of MT the information is missing for the first years, and then provided at a less aggregated level, but still usable. As is the case for MT, for the Dutch data are available at a more aggregated level, and the same applies to Denmark in 2014. Even though for the latter two countries the aggregation is not clearly documented, it seems that, primarily sizes 1-10 have be grouped into a single category (labeled "1").

Table 6. Number of categories used to report PL130 (maximum number is 15)

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
AT	15	15	15	15	15	15	15	15	15	15	15	15	14	15
CZ	0	15	15	15	15	15	15	14	14	14	14	14	14	14
DE	0	13	13	13	13	13	13	14	14	14	14	14	14	14
DK	13	13	13	13	13	13	13	13	13	13	4	13	13	13
FI	13	13	13	15	15	15	15	15	15	15	15	15	15	15
FR	14	14	14	14	14	14	14	14	14	14	14	14	14	14
IS	14	14	14	14	13	13	13	13	13	13	13	13	13	0
MT	0	0	0	15	6	6	6	6	6	6	6	6	6	6
NL	0	5	5	5	5	5	5	5	5	5	5	5	5	5
SE	13	13	15	15	15	15	15	15	15	15	15	15	15	15
SI	0	13	13	13	13	13	13	13	13	13	13	13	13	13
UK	0	11	11	11	13	13	13	13	15	15	15	15	15	15

In contrast to PL040 and PL050, the information is to be provided only for those *currently* working (as self-employed or employee), and not for the last job in case someone is currently inactive. This clearly creates some bias, as for employees all essential variables are in principle collected also for those who are no longer active. **Therefore, the GESIS routines to create social class switches all observations to missing if PL130 is missing, for both the self-employed and the employees. This implies that for everyone, social class is only available for those that define themselves as currently working** (first period: PL030= 1 or 2 or worked at least one hour in previous week; second period: PL031==1, 2, 3 or 4). Furthermore, countries working with the 'selected respondent model' (DK, FI, IS, NL, NO, SE, SI) have to provide this variable for the selected respondents only, even though SI has opted to provide this information both for selected respondents and other household members.

As a result, there are large differences in the number of self-employed (PL040==1 | PL040==2) for which this variable is available, and the number is very low for quite a few countries. Unfortunately, also within a non-negligible number of countries, there is quite some variation across time in the proportion of self-employed for which the variable is available. The worst performers are Denmark and Sweden, for which non-response is very high, with quite some cross-temporary variation. For Denmark, this variable is even missing completely between in SILC 2012-2015, implying that the self-employed cannot be categorised in ESeC for that period. Any long-term trend should consider this data-driven change in the composition of the higher social classes.

Table 7. Proportion of (former) self-employed (PL040==1 or PL040==2) at working age [18-65[in the sample for which PL130 is filled, countries with at least 5 percentage points between maximum and minimum value

selected resp.		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
	AT	0.78	0.83	0.80	0.27	0.28	0.76	0.76	0.78	0.78	0.80	0.80	0.78	0.81	0.84
	BE	0.80	0.82	0.81	0.84	0.85	0.87	0.87	0.85	0.87	0.84	0.88	0.84	0.84	0.86
	BG	.	.	.	1.00	0.90	0.88	0.90	0.87	0.83	0.84	0.85	0.84	0.84	0.82

selected resp.		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
	CH	.	.	.	0.94	0.95	0.91	0.88	0.82	0.89	0.83	0.83	0.86	0.79	.
	CY	.	0.85	0.82	0.81	0.84	0.81	0.82	0.79	0.79	0.71	0.75	0.71	0.73	0.77
	DE	.	0.78	0.84	0.92	0.91	0.89	0.82	0.82	0.82	0.82	0.81	0.84	0.83	0.85
S	DK	0.27	0.27	0.24	0.26	0.25	0.22	0.16	0.24	0.00	0.00	0.00	0.00	0.47	0.47
	EE	0.85	0.91	0.87	0.89	0.90	0.84	0.81	0.83	0.81	0.83	0.87	0.87	0.87	0.89
	EL	0.86	0.86	0.85	0.85	0.86	0.85	0.85	0.80	0.81	0.76	0.74	0.77	0.79	0.80
	ES	0.79	0.85	0.86	0.85	0.84	0.82	0.83	0.82	0.81	0.79	0.81	0.81	0.82	0.83
S	FI	0.54	0.45	0.44	0.44	0.45	0.40	0.41	0.42	0.42	0.42	0.42	0.42	0.42	0.42
	FR	0.80	0.84	0.82	0.78	0.74	0.78	0.77	0.77	0.80	0.80	0.76	0.76	0.75	0.77
	HR	.	.	.	.	.	.	0.78	0.73	0.76	0.70	0.71	0.72	0.72	0.70
	HU	.	0.80	0.74	0.75	0.76	0.78	0.73	0.76	0.75	0.75	0.73	0.71	0.81	0.80
	IE	0.91	0.95	0.90	0.89	0.87	0.83	0.79	0.77	0.76	0.77	0.87	0.93	0.86	0.86
S	IS	0.41	0.42	0.44	0.40	0.44	0.39	0.34	0.35	0.37	0.40	0.38	0.38	0.40	.
	IT	0.83	0.80	0.79	0.83	0.83	0.82	0.82	0.77	0.74	0.75	0.75	0.73	0.73	0.76
	LT	.	0.93	0.92	0.90	0.91	0.88	0.83	0.82	0.86	0.86	0.87	0.89	0.87	0.88
	LU	0.74	0.75	0.77	0.75	0.76	0.77	0.80	0.78	0.78	0.79	0.79	0.74	0.91	0.89
	LV	.	0.92	0.92	0.92	0.92	0.83	0.80	0.85	0.87	0.90	0.89	0.93	0.92	0.89
	MT	.	.	.	0.76	0.74	0.72	0.73	0.71	0.70	0.72	0.74	0.74	0.73	0.79
S	NO	0.45	0.49	0.51	0.49	0.48	0.48	0.43	0.46	0.43	0.43	0.50	0.46	0.41	0.47
	PT	0.85	0.83	0.83	0.83	0.82	0.75	0.74	0.75	0.71	0.67	0.73	0.74	0.76	0.77
	RO	.	.	.	0.93	0.92	0.76	0.85	0.86	0.89	0.88	0.90	0.90	0.90	0.91
	RS	.	.	.	.	.	.	.	.	.	0.67	0.80	0.75	0.74	0.76
S	SE	0.27	0.41	0.18	0.21	0.18	0.17	0.16	0.16	0.46	0.42	0.43	0.45	0.45	0.45
S	SI	.	0.64	0.75	0.74	0.78	0.78	0.78	0.76	0.72	0.71	0.72	0.73	0.79	0.79
	SK	.	0.90	0.90	0.89	0.92	0.90	0.90	0.88	0.86	0.85	0.84	0.85	0.87	0.89
	UK	.	0.86	0.82	0.83	0.86	0.81	0.82	0.83	0.84	0.86	0.87	0.89	0.89	0.88

3.2.4   Supervisory status (PL150)

In the case of employees, ESeC makes a distinction between those who supervise others and those who do not. In EU-SILC, this information is collected in a simply dummy variable for all employees and those currently inactive who were an employee in their last job. More in particular, the variable 'Managerial position' defines supervisory responsibility as 'the formal responsibility for supervising a group of other employees (other than apprentices), whom the respondent supervises directly, sometimes when doing some of the work that (s)he supervises.' In countries with the 'selected respondent model', this information has to be collected for selected respondents only.

Response rates seem to be somewhat better than those related to 'size of the local unit' are (mostly less than 3% item-non-response), with the huge exception of selected respondent countries, were response rates are below 50%. Unfortunately, also in this case there are several countries in which the response rate varies considerably across time. Even if item-non-response were completely at random, this affects the comparability of the middle class and upper class groups by artificially changing the share of

employees in these groups. It does not affect the composition of the lower classes in the same way, as these consist of employees only.

Table 8. Proportion of (former) employees (PL040==3) aged [18-65[in the sample for which information on managerial position (PL150) is available; countries with at least 5 p.p. difference between minimum and maximum value

		2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
	BE	0.70	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.99	0.99	0.99	1.00	0.99	1.00
	CH	.	.	.	0.98	0.98	0.99	0.98	0.99	0.93	0.99	0.99	0.99	0.96	.
	DE	.	0.72	0.73	0.86	0.92	0.93	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
S	DK	0.47	0.49	0.49	0.48	0.49	0.49	0.49	0.48	0.46	0.46	0.45	0.38	0.51	0.51
	EE	0.78	0.99	0.99	0.99	1.00	0.99	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99
	EL	0.90	0.91	0.91	0.92	0.93	0.94	0.94	0.94	0.94	0.95	0.95	0.96	0.96	0.96
S	FI	0.48	0.43	0.44	0.48	0.56	0.50	0.50	0.50	0.51	0.51	0.51	0.51	0.51	0.52
	FR	0.74	0.74	0.74	0.78	0.85	0.90	0.93	0.89	0.90	0.90	0.91	0.91	0.95	0.97
	HU	.	0.87	1.00	1.00	0.96	1.00	1.00	1.00	0.99	0.99	1.00	1.00	0.99	0.99
S	IS	0.40	0.40	0.40	0.40	0.46	0.44	0.45	0.44	0.44	0.43	0.43	0.43	0.43	.
S	NL	.	0.49	0.56	0.55	0.55	0.51	0.50	0.50	0.49	0.49	0.50	0.50	0.49	0.49
S	NO	0.48	0.49	0.54	0.55	0.55	0.53	0.49	0.48	0.46	0.47	0.51	0.45	0.45	0.45
	RO	.	.	.	0.91	0.91	0.92	0.94	0.94	0.96	0.94	0.94	0.95	0.95	0.95
	RS	.	.	.	.	.	.	.	.	.	0.92	0.95	0.97	0.98	0.98
S	SE	0.45	0.43	0.54	0.55	0.54	0.55	0.55	0.53	0.50	0.51	0.51	0.51	0.51	0.51
S	SI	.	0.44	0.41	0.41	0.41	0.36	0.37	0.37	0.38	0.38	0.38	0.38	0.39	0.38
	UK	.	1.00	1.00	0.98	0.97	0.97	1.00	1.00	0.80	0.81	0.82	0.82	0.82	0.98

Some caveats mentioned by NSIs in their report to Eurostat: In 2017 in Norway the question on managerial position has not been asked for the last job of the formerly employed and is therefore missing for them. In France, the way in which the question was asked and the response categories have changed from SILC 2011 onwards.

4 The variables combined

In this section, we put these variables together to construct the social class structure in accordance with the GESIS do-files. We have not carried out any additional corrections, except for those that required to cover the extended timespan 2004-2017 (for 2015-2017 no do-files were available at the time of writing) and the following minor changes: PL031 (self-defined current economic status) is only used from ESeC-08 onwards; ESeC-08 is switched to missing in case PL051 is equal to zero (only 16 cases in Ireland). This also implies that, mostly, ESeC is available for those currently at work (self-defined economic status) and not for those out of work but previously in work, although there are some exceptions to this rule until EU-SILC 2008 (incl.).

4.1 The transition from ISCO-88 to ISCO-08: ESeC-88 vs ESeC-08

In the table below we compare for EU-SILC 2011 both versions of social class, showing the relevance of taking this break in series into account.

Table 9. Unweighted percentage of the total sample which moves at least one or at least two classes up or down in a 3 class schema when comparing ESeC-88 with ESeC-08 in EU-SILC 2011 (only part of sample for which both class variables are available)

	at least 1	at least 2
AT	13.8	1.4
BE	10.6	0.6
CH	16.4	1.0
CY	12.0	0.1
CZ	13.2	0.7
DE	19.4	0.9
DK	27.5	1.0
EE	9.7	0.7
ES	11.5	1.3
FR	12.4	0.7
GR	4.3	0.1
HR	12.8	0.4
HU	9.8	0.4
IS	5.3	0.1
IT	16.5	2.6
LT	6.0	0.4
LU	12.3	0.1
LV	12.6	1.7
MT	7.6	0.1
NL	15.1	2.3
NO	22.4	3.1
PL	8.4	0.3
PT	10.7	0.4
RO	5.0	0.1
SE	15.5	0.4
SI	12.4	0.5
SK	20.8	1.5
UK	15.4	0.7
Total	13.0	0.9

Table 10. Cross-tabulation (unweighted) of ESeC-88 and ESeC-08 collapsed into three categories, EU-SILC 2011

ESeC08	ESeC88			
	1	2	3	Total
1	32.11	7.4	0.45	39.96
2	1.67	20.95	0.14	22.76
3	0.47	2.81	33.99	37.28
Total	34.25	31.16	34.59	100

Although, broadly speaking, the overlap between both class variables is considerable, about 13 per cent of the sample is classified in a different social class in ESeC-08 as compared to ESeC-88, while slightly less than 1 per cent of the sample moves two classes up or down. This average score for the entire EU-SILC 2011 sample conceals considerable variation in the overlap between both variables across countries. In nine countries more than 15 per cent of the sample moves at least one class up or down, while in Germany, Denmark, Norway and Slovakia this is the case for close to or more than 20 per cent of the sample. Only in Greece and Romania, this number is below five per cent. In other words, for most countries the change from ESeC-88 to ESeC-08 is a substantial change in defining social class, implying a break in series that cannot easily be ignored. In principle, it should be possible to develop a crosswalk between ISCO-08 and ISCO-88, but give the relatively high level of aggregation of the data in EU-SILC, we did not undertake such an attempt.

4.2   Social class: only for those currently in work

In theory, one's social class does not immediately change when someone becomes unemployed or (early) retired. In line with this, the GESIS procedure starts from variable PL040 to define one's economic status (employee vs. self-employed), regardless of whether or not one is currently at work. However, one of the other variables that is used for constructing social class (notably size of the local unit) is collected only for those currently at work (as defined by their self-defined current activity status, PL030 / PL031). Therefore, for everyone (and not just the self-employed) the social class variable is switched to missing when PL130 (size of local unit) is not available. Yet, this does not solve the problem entirely, as can be seen from the table below. Since EU-SILC 2009, all those for whom the social class variable is defined, are currently at work. In contrast, this is not in all countries the case for the period up to 2008 (incl.) in which for some countries more than 5 per cent are currently out of work, amounting to close to ten or more per cent for some years in Switzerland, Iceland, Norway and the United Kingdom. As far as the latter country is concerned, though, this is largely (but not fully) the result of missing values on PL030. However, in the other countries excluding missing values on PL030 does not change the picture. In any case, to depict a consistent trend of what happened to the relation between social class and other variables, it seems the best strategy to limit the analysis to those currently in work. Therefore, we recommend to either restrict the analysis explicitly to this subpopulation, or to switch the social class variable to zero for those who do not indicate in PL030 or PL031 to be currently at work. If one could expect that non-response patterns for the variable on the size of the local economic unit (PL130) would differ between the self-employed and employees, one could even opt for switching social class to missing when PL130 is missing only in the case of the self-employed, and applying the 'currently at work' condition to both the self-employed and employees.

Table 11. Unweighted proportion of those currently at work (variable PL030 / PL031) among those for whom the social class variable is defined (working age population only)

	2004	2005	2006	2007	2008	2009	2010	2011
CH				0.89	0.90	1	1	1
DE		1	1	0.95	0.95	1	1	1
FR	0.98	0.98	0.98	0.97	1	1	1	1
HU		1	0.94	1	1	1	1	1
IS	1	1	0.92	0.90	0.91	1	1	1
IT	0.95	0.98	0.98	0.99	0.99	1	1	1
NO	0.92	0.96	0.93	1	1	1	1	1
PL		0.96	0.97	0.97	0.97	1	1	1
RO				0.96	0.96	1	1	1
UK		0.88	0.90	0.91	0.88	1	1	1

4.3   Missing values for social class over time

If we restrict the sample to those at working age (aged [18-65[) and those who say they are in work (PL030<3 or PL031 <5), the proportion of the sample for whom social class is avialable, is typically above 95 per cent. However, it is much lower in 'selected respondent countries', and for some countries there is substantial variation across time, especially in the first period of observation (2004-2011). For countries with the selected respondent model this gives a somewhat distorted picture, as, in principle, non-response to a good degree has been implemented deliberately, and randomly. Therefore, we reproduced the same tables, but now limiting the sample to selected respondents. However, even in these cases data availability is somewhat lower for most of these countries.

Table 12. Proportion of the unweighted sample at working age and currently in work (PL030<3 or PL031<5) for which ESeC-88 is defined

	2004	2005	2006	2007	2008	2009	2010	2011
AT	0.98	0.98	0.99	0.92	0.91	0.96	0.99	0.99
BE	0.94	0.97	0.99	0.98	0.98	0.96	0.97	0.93
BG	.	.	.	0.95	0.98	0.99	0.99	0.00
CH	.	.	.	0.99	0.99	0.99	0.99	0.99
CY	.	0.99	0.99	0.99	0.99	1.00	1.00	1.00
CZ	.	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DE	.	0.98	0.96	0.93	0.98	0.98	0.99	0.99
DK	0.40	0.43	0.38	0.45	0.44	0.43	0.42	0.43
EE	0.98	0.99	0.99	0.99	0.99	0.99	0.99	0.98
EL	0.93	0.94	0.94	0.95	0.95	0.96	0.96	0.96
ES	0.95	0.94	0.93	0.94	0.95	0.94	0.94	0.94
FI	0.15	0.46	0.45	0.48	0.49	0.49	0.49	0.00
FR	0.97	0.96	0.96	0.95	0.87	0.92	0.88	0.83
HR	.	.	.	.	.	.	0.97	0.94

	2004	2005	2006	2007	2008	2009	2010	2011
HU	.	0.81	1.00	0.99	0.97	0.99	0.99	0.99
IE	0.96	0.96	0.96	0.96	0.96	0.97	0.96	0.00
IS	0.43	0.43	0.42	0.41	0.44	0.44	0.45	0.43
IT	0.98	0.97	0.98	0.99	0.98	0.99	0.98	0.97
LT	.	0.98	0.97	0.98	0.97	0.97	0.98	0.97
LU	0.98	0.98	0.98	0.98	0.99	0.98	0.98	0.98
LV	.	0.98	0.99	0.99	0.99	0.98	0.99	0.99
MT	.	.	.	0.98	0.98	0.98	0.98	0.99
NL	.	0.42	0.46	0.46	0.46	0.47	0.47	0.46
NO	0.47	0.47	0.51	0.52	0.51	0.51	0.50	0.47
PL	.	0.96	0.96	0.96	0.96	0.97	0.97	0.97
PT	0.99	0.99	0.99	0.98	0.99	0.97	0.97	0.98
RO	.	.	.	0.92	0.92	0.93	0.94	0.94
RS	.	.	.	.	.	.	.	.
SE	0.45	0.44	0.45	0.47	0.46	0.46	0.46	0.45
SI	.	0.30	0.32	0.33	0.33	0.34	0.35	0.35
SK	.	0.98	0.98	0.99	0.99	0.99	1.00	0.99
UK	.	0.98	0.97	0.96	0.99	0.94	0.94	0.92

Note: Shaded cells indicate a drop or increase of more than 5 percentage points

Table 13. Proportion of the unweighted sample at working age and currently in work (PL030<3 or PL031<5) for which ESeC-08 is defined

	2010	2011	2012	2013	2014	2015	2016	2017
AT	0.00	0.99	0.96	0.99	0.99	0.99	0.99	0.99
BE	0.00	0.94	0.98	0.97	0.98	0.98	0.98	0.99
BG	0.00	1.00	0.99	1.00	0.99	0.99	0.99	0.99
CH	0.99	0.99	0.99	0.99	0.99	0.99	0.95	.
CY	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
CZ	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
DE	0.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99
DK	0.00	0.44	0.43	0.44	0.36	0.42	0.48	0.48
EE	0.99	0.98	0.98	0.98	0.97	0.98	0.99	0.99
EL	0.00	0.96	0.97	0.97	0.97	0.98	0.97	0.97
ES	0.94	0.94	0.96	0.94	0.96	0.95	0.95	0.96
FI	0.00	0.48	0.49	0.50	0.50	0.49	0.49	0.50
FR	0.00	0.81	0.90	0.90	0.86	0.87	0.87	0.87
HR	0.00	0.94	0.96	0.99	0.99	0.99	0.98	0.93
HU	0.00	0.99	0.99	0.97	0.99	0.99	0.98	0.89
IE	0.00	0.97	0.96	0.96	0.96	0.95	0.94	0.95
IS	0.45	0.43	0.43	0.43	0.00	0.00	0.00	.
IT	0.00	0.97	0.97	0.98	0.98	0.98	0.99	0.99
LT	0.00	0.97	0.98	0.98	0.98	0.98	0.98	0.98

	2010	2011	2012	2013	2014	2015	2016	2017
LU	0.00	0.98	0.98	0.97	0.98	0.97	0.99	0.99
LV	0.00	0.99	0.99	0.99	1.00	1.00	1.00	0.98
MT	0.00	0.99	0.98	0.99	0.99	0.99	0.99	0.99
NL	0.00	0.46	0.46	0.46	0.47	0.48	0.49	0.49
NO	0.00	0.45	0.48	0.50	0.49	0.49	0.49	0.50
PL	0.00	0.97	0.98	0.98	0.97	0.98	0.97	0.97
PT	0.00	0.98	0.99	0.99	0.99	0.99	0.99	0.99
RO	0.00	0.94	0.96	0.94	0.95	0.95	0.96	0.96
RS	.	.	.	0.93	0.93	0.96	0.96	0.97
SE	0.00	0.45	0.47	0.48	0.49	0.48	0.50	0.50
SI	0.00	0.35	0.35	0.35	0.35	0.36	0.36	0.36
SK	0.00	0.99	0.99	1.00	1.00	1.00	1.00	1.00
UK	0.00	0.93	0.97	0.98	1.00	0.99	0.99	0.97

Note: Shaded cells indicate a drop or increase of more than 5 percentage points.

Table 14. Proportion of the unweighted sample *of selected respondents* at working age and currently in work (PL030<3 or PL031<5) for which ESeC-88 / ESeC-08 is defined, 'selected respondent countries'

ESeC-88								
	2004	2005	2006	2007	2008	2009	2010	2011
DK	0.81	0.86	0.78	0.93	0.90	0.89	0.88	0.92
FI	0.30	0.94	0.94	1.00	1.00	1.00	0.99	0.00
IS	0.96	0.97	0.96	0.93	0.98	0.98	0.98	0.98
NL	.	0.82	0.91	0.90	0.90	0.90	0.90	0.89
NO	0.94	0.94	0.98	0.98	0.98	0.98	0.96	0.92
SE	0.91	0.87	0.88	0.93	0.92	0.91	0.91	0.91
SI	.	0.90	0.94	0.94	0.95	0.95	0.96	0.96
ESeC-08								
	2010	2011	2012	2013	2014	2015	2016	2017
DK		0.94	0.91	0.91	0.75	0.85	0.94	0.91
FI		0.99	0.99	1.00	1.00	1.00	1.00	1.00
IS	0.98	0.98	0.99	0.98	0.00	0.00	0.00	.
NL		0.90	0.90	0.91	0.91	0.92	0.94	0.93
NO		0.87	0.94	0.97	0.97	0.97	0.96	0.97
SE		0.91	0.97	0.98	0.99	0.97	0.99	0.99
SI		0.96	0.95	0.95	0.95	0.95	0.95	0.95

4.3.1   Bias in the case of effective non-response

The main question is, though, whether these variations in data availability affect comparability over time and between countries. Whether or not this is the case, depends very much on the research question. In what follows, we focus on bias in terms of earnings. In the case of earnings, missing data seems to affect

both average earnings and earnings inequality. We computed a very simple measure of earnings consisting of cash, near-cash and non-cash employee income as well as cash profits or losses from self-employment (preferably gross, but when gross amounts were lacking, we took net amounts if available). Earnings were bottom-coded at zero EUR (non bottom-coded results are similar with regard to the main conclusions), and did not apply any top-coding. This earnings variable in itself is not without its problems for comparative research, but given that here we focus on the difference between cases for which ESeC is available and cases for which it is not, that should not be too big a problem. More precisely, for the countries with relatively big changes in the availability of the social class variable, we checked whether average earnings and earnings inequality among those for whom ESeC-88 is available is very different from the average earnings and earnings inequality among those for whom the ESeC-88 variable is not available. We restricted the sample to those currently at work (PL030<3 or PL031<5), but did not apply an age restriction. We weighted the sample with the appropriate cross-sectional survey weights (PB040).

Table 15. Ratio of average earnings among those in work for whom ESeC-88 is not available and those in work for whom ESeC-88 is available

country	year	Estimate	SE	LB	UB	DF	Response rate
AT	2007	0.86	0.04	0.78	0.93	6805	0.92
	2008	0.82	0.04	0.73	0.90	5706	0.91
	2009	0.90	0.07	0.77	1.02	5875	0.96
	2010	0.61	0.10	0.41	0.80	6187	0.99
	2011	0.54	0.08	0.37	0.70	6186	0.99
DE	2007	0.35	0.02	0.32	0.38	14152	0.93
	2008	0.66	0.07	0.53	0.79	13311	0.98
	2009	0.48	0.06	0.37	0.60	13086	0.98
	2010	0.43	0.05	0.33	0.53	13078	0.99
	2011	0.68	0.12	0.45	0.91	13511	0.99
FR	2007	0.62	0.03	0.57	0.67	10497	0.95
	2008	0.51	0.02	0.47	0.55	10417	0.87
	2009	0.67	0.02	0.63	0.71	10602	0.92
	2010	0.84	0.03	0.79	0.89	11043	0.88
	2011	0.92	0.03	0.86	0.98	11359	0.83
UK	2007	0.87	0.05	0.78	0.96	9274	O.96
	2008	0.70	0.08	0.54	0.86	8934	0.99
	2009	0.73	0.04	0.65	0.80	8361	0.94
	2010	0.74	0.05	0.65	0.84	8108	0.94
	2011	0.71	0.03	0.64	0.77	8057	0.92

Note: The response rate refers to the share of the sample at working age and currently in work for which ESeC-08 is defined.

The table above shows that, on average, earnings are considerably lower for those in work for whom ESeC-88 is not available as compared to the average earnings for whom ESeC-88 is available. Typically, this will not affect very much overall average earnings, given that in many cases ESeC-88 is missing for not much more than 10 per cent of the sample. Probably the most extreme case in the set is Germany 2007, where average earnings are more than 60 per cent lower for those with missing values on ESeC-88

(account for seven per cent of those in work in the sample). In this (rather extreme) case, overall average earnings for those in work would be over-estimated by six per cent if we would limit the sample to those with observations on ESeC-88. In contrast, in the UK 2010 average earnings for those in work would be overestimated by slightly less than two per cent if the sample would be restricted to those with an observed social class.

While average earnings tend to be lower for missing observations, earnings inequality among them tends to be (much) higher, as is demonstrated by the table below. As a result, overall earnings inequality that we observe for the sample with observations for social class tends earnings inequality in the total population in particular countries and years.

Table 16. The Gini coefficient of earnings for those in work for whom ESeC-88 is not available minus the Gini coefficient of earnings for those in work for whom ESeC-88 is available

country	year	Difference sample without ESeC-88 minus sample with ESeC-88	SE	LB	UB	DF	Gini for observed ESeC-88	Response rate
AT	2007	0.08	0.02	0.05	0.11	6805	0.38	0.92
	2008	0.04	0.02	-0.01	0.08	5706	0.41	0.91
	2009	0.06	0.02	0.01	0.10	5875	0.41	0.96
	2010	0.07	0.05	-0.02	0.17	6187	0.42	0.99
	2011	0.22	0.05	0.12	0.31	6186	0.40	0.99
DE	2007	0.15	0.01	0.12	0.17	14152	0.39	0.93
	2008	0.17	0.03	0.11	0.23	13311	0.42	0.98
	2009	0.23	0.04	0.16	0.30	13086	0.41	0.98
	2010	0.20	0.03	0.15	0.25	13078	0.41	0.99
	2011	0.21	0.05	0.11	0.31	13511	0.41	0.99
FR	2007	0.10	0.02	0.06	0.13	10497	0.36	0.95
	2008	0.15	0.02	0.12	0.18	10417	0.34	0.87
	2009	0.02	0.01	0.00	0.05	10602	0.36	0.92
	2010	0.02	0.02	-0.01	0.05	11043	0.36	0.88
	2011	0.03	0.02	0.00	0.07	11359	0.37	0.83
UK	2007	0.04	0.02	0.00	0.08	9274	0.42	O.96
	2008	0.14	0.04	0.06	0.21	8934	0.46	0.99
	2009	0.06	0.02	0.02	0.10	8361	0.45	0.94
	2010	0.10	0.03	0.05	0.15	8108	0.45	0.94
	2011	0.02	0.02	-0.02	0.05	8057	0.46	0.92

Note: The response rate refers to the share of the sample at working age and currently in work for which ESeC-08 is defined.

How may this affect a decomposition of income inequality by social class? This may be affected in two ways: it may affect the level of earnings in particular classes and the share of social classes. The implications are different for each type of bias. If non-response patterns are the result of random non-response in particular classes, then this will not necessarily affect the estimate of between-class inequality

as judged from, for instance, the ratio of median incomes by social class. However, by changing the estimated population share of particular classes, it still may have an impact on the results of a within vs. between group inequality decomposition. So it is important to look both at the results of decompositions and descriptive measures of median/average incomes for individual classes. If, in contrast, non-response affects also the distribution of income within some, or all, classes, both trends within and between classes will be affected in more unpredictable ways.

4.3.2 Bias in the cases of countries with the 'selected respondent model'

In the case of selected respondent countries, non-response occurs deliberately and, if all goes well, randomly. Of course, also in this case there may still be some non-deliberate non-response, but we would expect it to have a much more moderate effect than what the overall low response rate would suggest. The graphs below depict the ratio (as a percentage) of the average income of those without a value on social class (including non-selected respondents) and the average income for whom we can observe social class and the difference in the Gini coefficient.

Figure 3. The ratio of average earnings and the difference in Gini coefficients between those in work without an observed ESeC class as compared to those in work with an observed ESeC class in countries with the selected respondent model, EU-SILC

NO

SE

SI
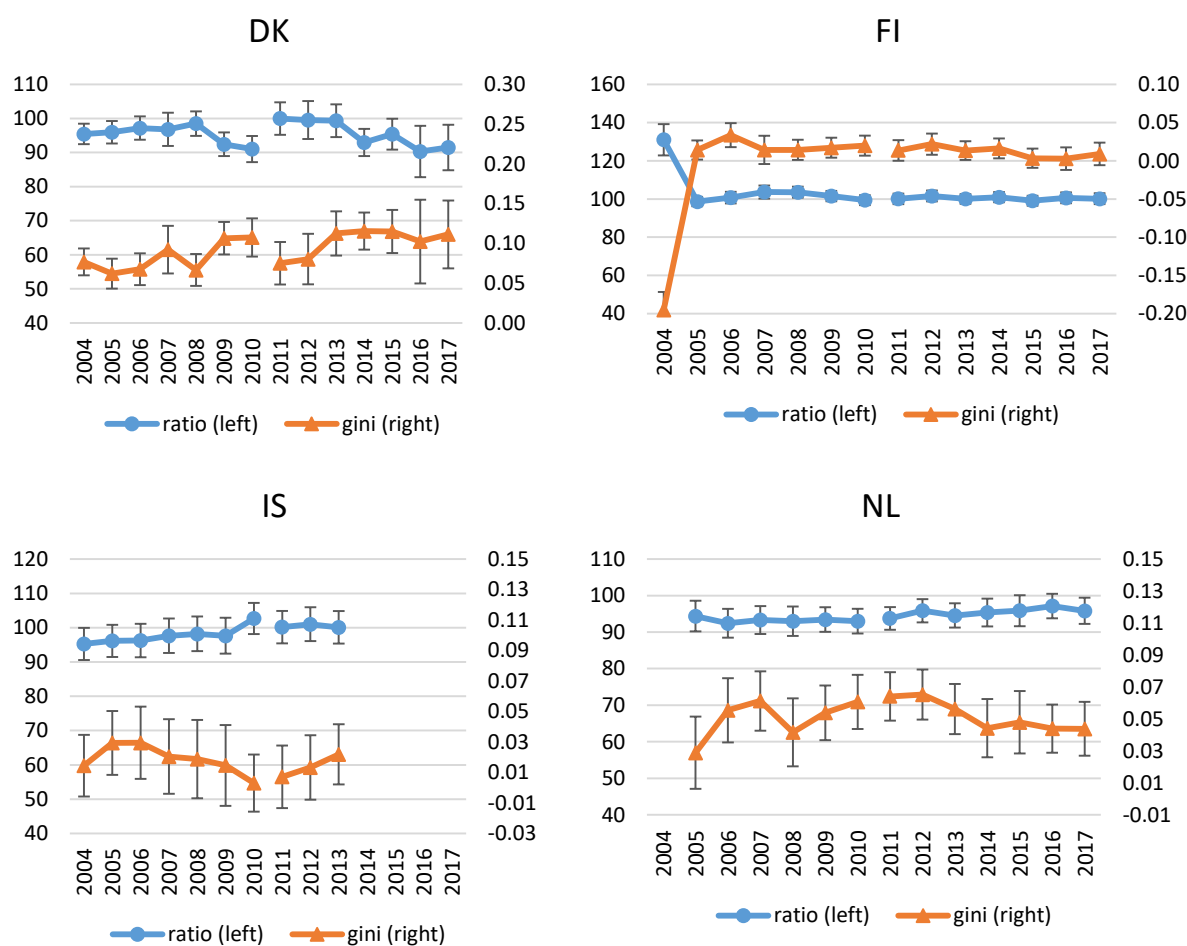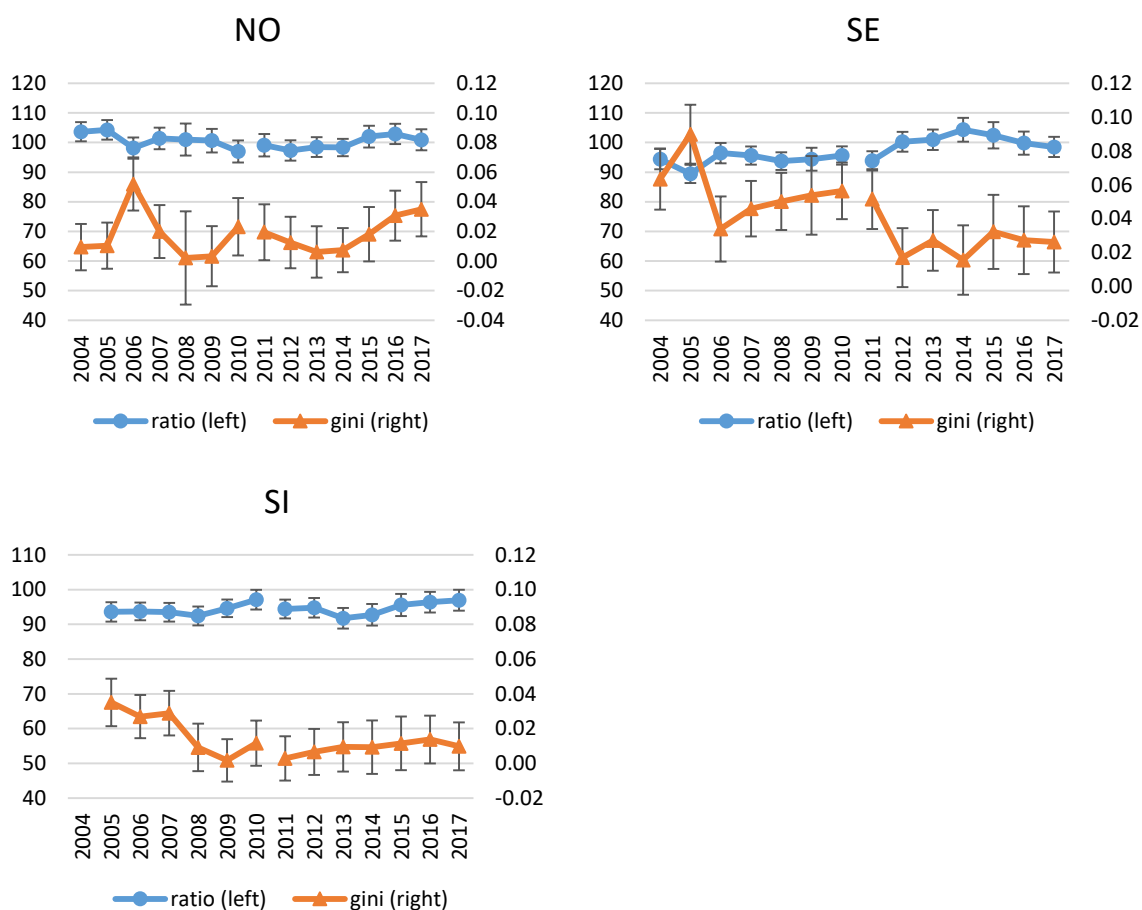
Notes: ESeC-88 until 2010 (incl.), ESeC-08 thereafter. 95% confidence intervals.

What can we learn from this? (1) In comparison with the other countries, the difference in average earnings is much lower, and those with missing values on ESeC typically have average incomes above 90% and in several countries close to 100%. (2) With the exception of Denmark, the Netherlands and Sweden at some point, the difference in earnings inequality tends to be relatively low. In other words, the problem of selective non-response seems to be less severe, on average. (3) In some countries bias tends to vary quite a bit over time, as was also the case for the other countries. In conclusion, the number of observations will be lower in the selected respondent countries, but if one pays attention to changes in response rates and bias over time, the data do not seem to be of much lower quality than that of some of the other countries.

This being said, we should take into account that there are two types of missing data at work here: deliberate, in principle random, missing observations due to non-selection, and non-random missing observations due to non-response. The previous graphs tell something about the combination of this 'random' and 'deliberate' non-response. To assess the extent of more problematic non-response, we also redid the same analysis, but now limiting the sample to 'selected respondents' only. As previously, but now in a much more pronounced way, average earnings among non-respondents tend to be lower, but go hand in hand with higher levels of inequality. Several country-specific elements are worth noting. For instance, Finland displays a remarkably low response rate in 2004, with non-respondents having higher

average earnings and much lower inequality, indicating that clearly high earners are missing among the respondents. The increase in response rates to nearly 100% from 2007 onwards is very reassuring on the other hand, but one has to be careful with the changes between 2005 and 2007. Also for the Netherlands, one should be careful with the change between 2005 and 2006, and in Norway the change between 2011 and subsequent years. Obviously, it is not just changes over time that matter, but cross-country differences as well. It seems that that some (especially lower) earnings categories are missing. Given the variation in cross-country response rates, if these earnings categories are concentrated in particular classes, this will not just affect within country trends in within vs between class inequalities, but also patterns across countries.

Figure 4. The ratio of average earnings and difference in Gini coefficients between those in work without an observed ESeC class as compared to those in work with an observed ESeC class, selected respondents only, EU-SILC

ratio of average earnings in IS

Difference in Gini earnings in IS

ratio of average earnings in NL

Difference in Gini earnings in NL

ratio of average earnings in NO

Difference in Gini earnings in NO
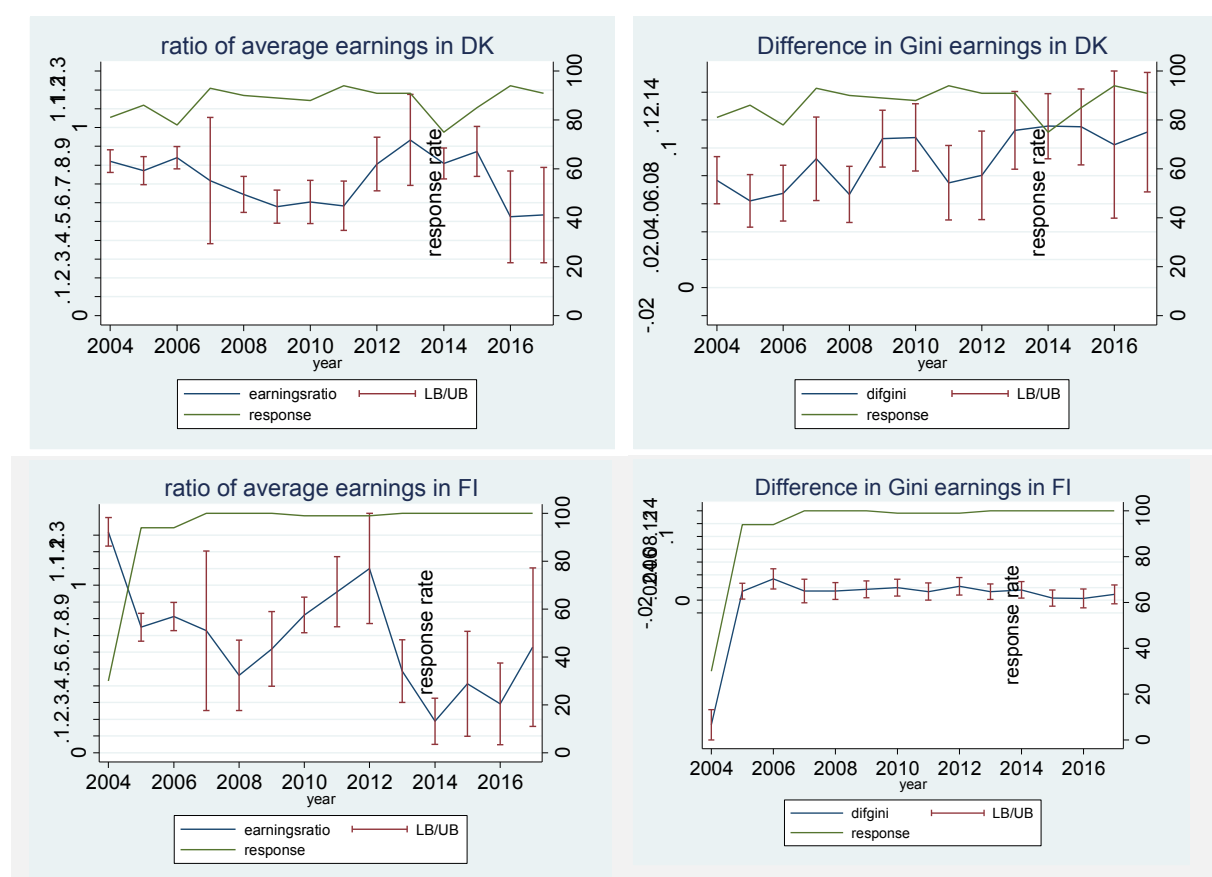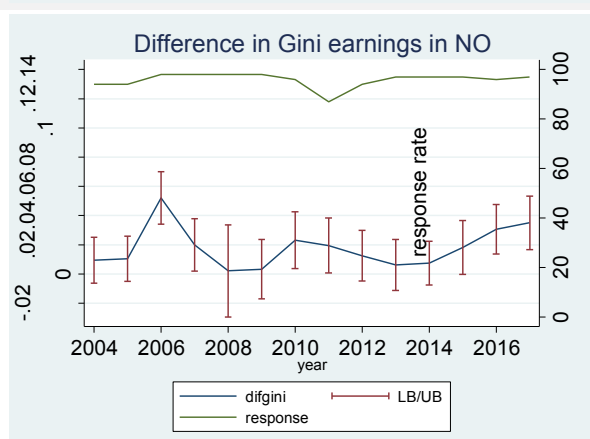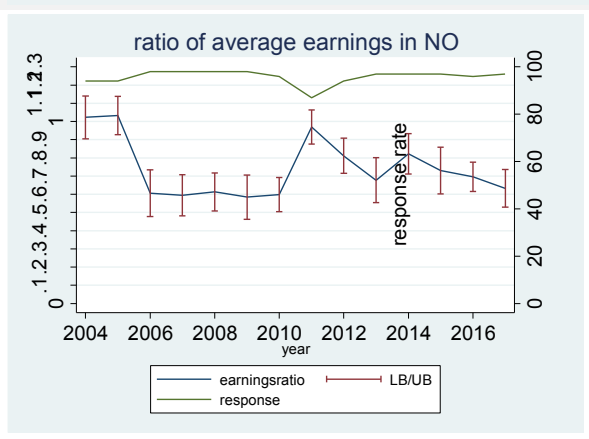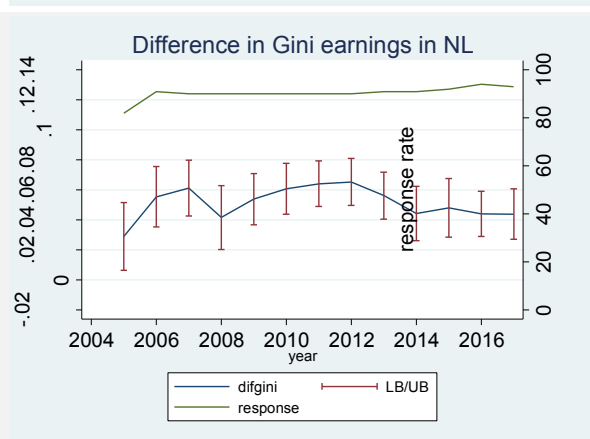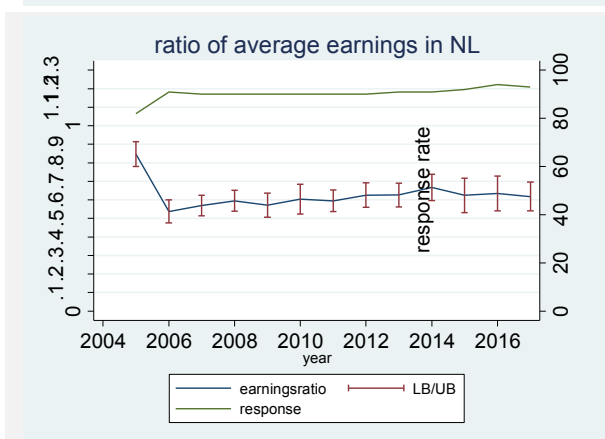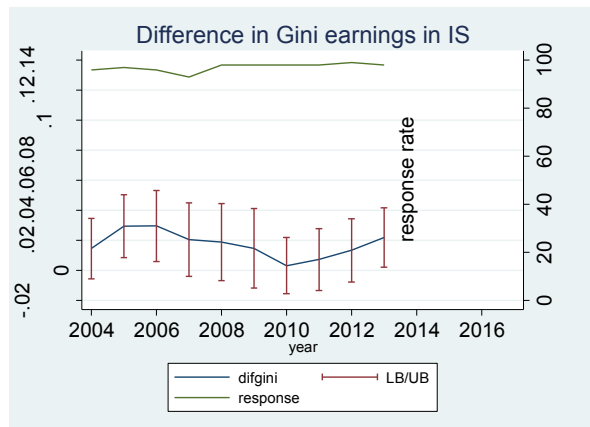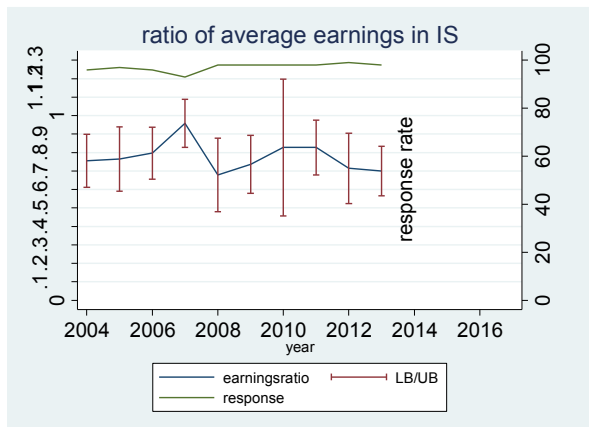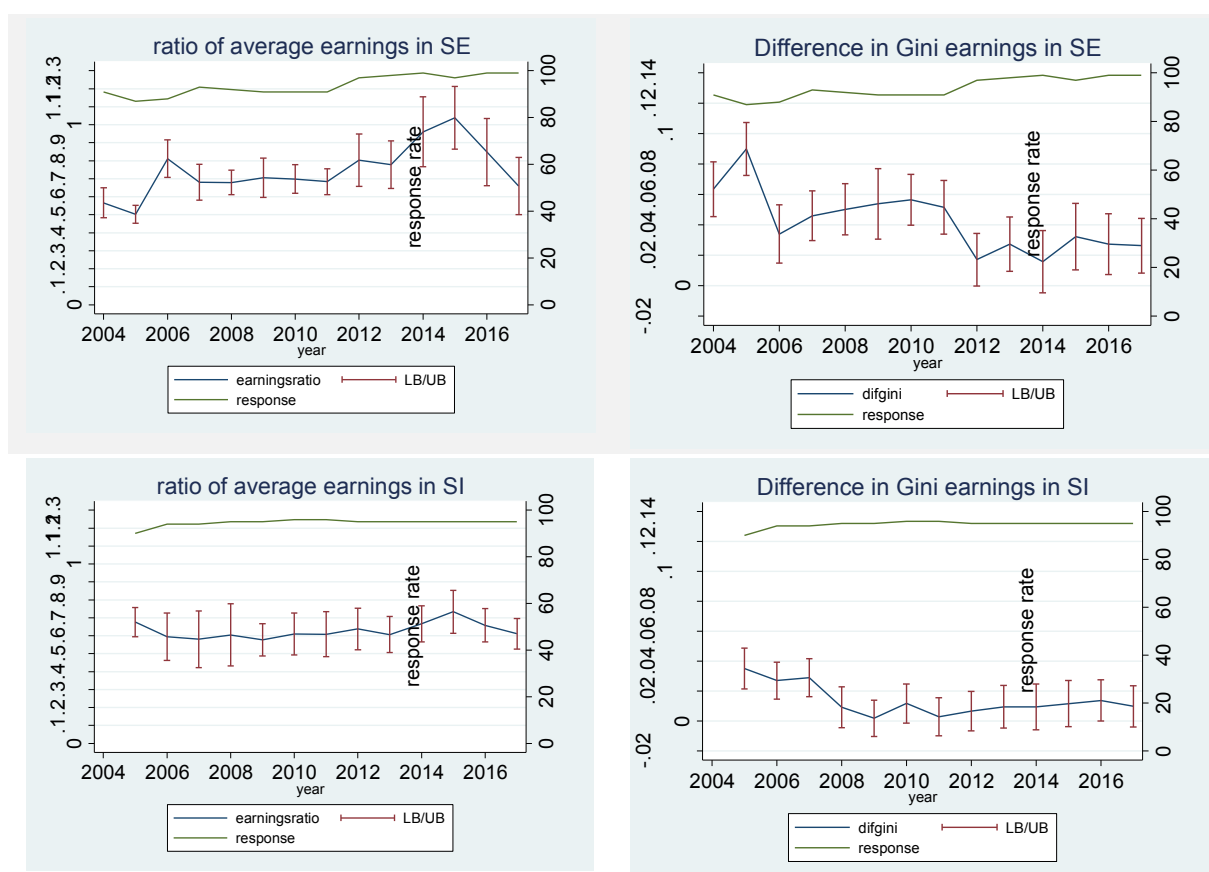
Notes: ESeC-88 until 2010 (incl.), ESeC-08 thereafter. 95% Confidence intervals.

4.4   We should be careful with the share of self-employed

In the previous parts, we pointed out that one should be careful with data availability with regard to the size of the economic unit (PL130) and the supervisory role of employees (PL150). Missing data on these variables results in missing information on ESeC. In turn, this may affect the relative share of self-employed and employees in the middle and upper classes. It goes without saying that for many types of analysis representativeness may be threatened by non-negligible variations in the share of the self-employed, if caused by data quality issues rather than real trends in society. To simplify the analysis that follows, we regrouped the original nine ESeC classes into three classes: the working class (no self-employed), the middle class (ESeC class 3 to 6 (incl.), the upper class (ESeC class 1 and 2).

From the tables that follow, it appears that there are quite a few countries with remarkable changes in the share of self-employed in the upper and middle class. For instance, in Finland in 2004, the upper two classes consist exclusively of self-employed, while in Denmark the self-employed are completely out of the picture from 2012 until 2015 (incl.). In Austria, in contrast, the self-employed account for a strongly reduced share in 2007 and 2008. In contrast, in Hungary in 2010 the share of self-employed in the upper class increases from 0.12 to 0.23. The tables below that these sudden drops or increases in the share of self-employed occur also in other countries.

One solution might be to focus on employees only. However, the share of the self-employed varies not only across time, but also across countries. While in some countries (EE, SE, SI) the share of the self-employed in the upper class is about 5 per cent, it is above 20 per cent in some others (GR, IT). The average share of and variation in the share of the self-employed in the middle class is even much higher. If one is interested in studying how social classes are doing, restricting oneself to the employees implies leaving out a (strongly) varying share of those classes in different countries. Finally, it is worth mentioning that for some countries the share of self-employed marks a shift in 2011 when going from ESeC-88 to ESeC-08. Sometimes this is the reflection of a longer trend, but for some countries, it marks a noticeable difference between the two periods.

Table 17. Weighted proportion of self-employed (PL040 < 3) in the upper class at working age, *ESeC-88*, countries with fluctuation of at least 0.05 points (working age population only)

	2004	2005	2006	2007	2008	2009	2010	2011
AT	0.10	0.08	0.07	0.04	0.04	0.08	0.08	0.10
BG	.	.	.	0.05	0.16	0.14	0.13	.
DE	.	0.13	0.11	0.12	0.09	0.08	0.07	0.06
EL	0.20	0.18	0.20	0.22	0.23	0.23	0.24	0.22
FI	1.00	0.07	0.07	0.08	0.07	0.06	0.07	.
HU	.	0.13	0.13	0.10	0.09	0.12	0.23	0.19
LV	.	0.06	0.08	0.08	0.11	0.07	0.09	0.07
MT	.	.	.	0.09	0.05	0.04	0.05	0.04
PT	0.15	0.12	0.11	0.15	0.13	0.12	0.10	0.11
RO	.	.	.	0.17	0.14	0.12	0.13	0.11
SE	0.08	0.09	0.04	0.05	0.04	0.04	0.02	0.04

Table 18. Weighted proportion of self-employed (PL040 < 3) in the upper class at working age, *ESeC-08*, countries with fluctuation of at least 0.05 points (working age population only)

	2010	2011	2012	2013	2014	2015	2016	2017
DK	.	0.06	0.00	0.00	0.00	0.00	0.04	0.06
EL	.	0.22	0.26	0.26	0.25	0.23	0.22	0.21
HU	.	0.18	0.16	0.16	0.09	0.10	0.10	0.11
IS	0.06	0.07	0.08	0.13	.	.	.	.
SE	.	0.03	0.08	0.08	0.08	0.09	0.08	0.08
SK	.	0.12	0.16	0.17	0.17	0.17	0.18	0.18

Table 19. Share of self-employed (PL040 < 3) in the middle class at working age, *ESeC-88*, countries with fluctuation of at least 0.05 points (working age population only)

	2004	2005	2006	2007	2008	2009	2010	2011
AT	0.31	0.25	0.27	0.10	0.10	0.23	0.25	0.25
BE	0.25	0.24	0.24	0.23	0.21	0.22	0.22	0.20
BG	.	.	.	0.35	0.45	0.43	0.43	.
CY	.	0.31	0.29	0.31	0.30	0.29	0.27	0.24
DK	0.14	0.12	0.11	0.09	0.06	0.09	0.07	0.09
EE	0.25	0.26	0.27	0.25	0.23	0.29	0.27	0.27
EL	0.63	0.59	0.59	0.61	0.60	0.58	0.57	0.60
FI	1.00	0.39	0.39	0.38	0.37	0.38	0.37	.
IS	0.31	0.32	0.32	0.28	0.31	0.31	0.28	0.25
LT	.	0.49	0.41	0.36	0.38	0.37	0.36	0.39
LV	.	0.30	0.25	0.23	0.24	0.26	0.28	0.27
PT	0.50	0.48	0.47	0.47	0.42	0.39	0.40	0.37
RO	.	.	.	0.62	0.62	0.58	0.63	0.68
SE	0.12	0.22	0.09	0.11	0.11	0.10	0.09	0.09
SI	.	0.16	0.16	0.14	0.20	0.21	0.20	0.22

Table 20. Share of self-employed (PL040 < 3) in the middle class at working age, *ESeC-08*, countries with fluctuation of at least 0.05 points (working age population only)

	2010	2011	2012	2013	2014	2015	2016	2017
DE	.	0.14	0.13	0.11	0.11	0.09	0.09	0.08
DK	.	0.18	0.00	0.00	0.00	0.00	0.09	0.08
EE	0.40	0.39	0.36	0.33	0.38	0.38	0.39	0.38
HU	.	0.47	0.45	0.42	0.42	0.40	0.40	0.38
IE	.	0.39	0.37	0.36	0.36	0.36	0.35	0.34
LT	.	0.54	0.56	0.56	0.61	0.57	0.58	0.51
LU	.	0.19	0.19	0.20	0.19	0.19	0.15	0.14
LV	.	0.37	0.39	0.40	0.34	0.41	0.46	0.44
NL	.	0.31	0.35	0.32	0.30	0.31	0.30	0.29
NO	.	0.27	0.26	0.25	0.26	0.28	0.24	0.23
PT	.	0.47	0.42	0.38	0.39	0.36	0.38	0.37
RS	.	.	.	0.66	0.59	0.52	0.50	0.49
SE	.	0.16	0.38	0.34	0.34	0.32	0.32	0.37
SI	.	0.25	0.26	0.27	0.21	0.22	0.20	0.23
SK	.	0.46	0.54	0.52	0.50	0.49	0.47	0.44
UK	.	0.38	0.40	0.42	0.42	0.43	0.41	0.44

5   An alternative coding of the self-employed by size of firm

In the code published on the GESIS website, first the self-employed are identified as a group, and subsequently they are subdivided into two categories: those in a firm with over 10 employees, and others. In EU-SILC, the variable on the employment status allows users to make a difference between self-employed with employees and self-employed without employees. Given that the logic of the distinction by size of the firm is to make a difference between large and small employers, it would be logical to apply this criterion only to those persons who indicate in PL040 that they are a self-employed with employees. In the GESIS code, however, this information is left unused. Consequently, some self-employed who do not indicate to have employees, are categorised as large employers. In what follows, I show that making full use of the information contained in PL040 is probably worthwhile the effort.

Figure 5. Treatment of self-employed in GESIS code: self-employed without employees can be classified as 'large employer'
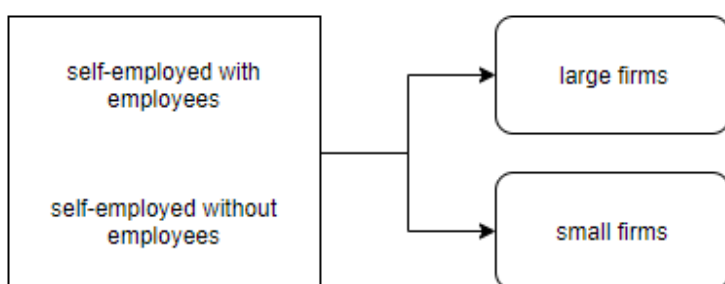
Figure 6. Alternative proposal: self-employed without employees are always classified as 'small employer'
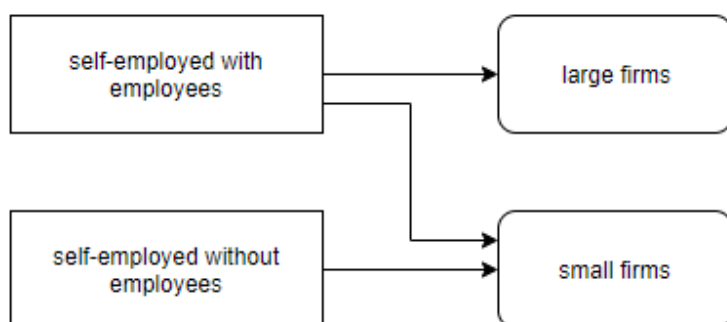
This may seem a minor difference, as one would expect that very few self-employed who indicate to have no employees would situate themselves in a larger firm. However, depending on the country and time period, this situation seems at times not to be very uncommon. Taking the entire sample together (EU-SILC 2004-2017, excl. MT), 46% of those who were a self-employed with many employees according to the old code, are now classified as self-employed with few or no employees. Yet, ESeC does classify some 'small self-employed' still in the upper class, so the impact on the ESeC classification is somewhat less dramatic: less than 20 per cent of the self-employed, working respondents at active age see their ESeC class change as a consequence of the change in coding. The impact of the change is illustrated in more detail in the tables below.

Table 21. Number of self-employed respondents that change social class (ESeC) in the new code, persons at active age and in work only, all countries except Malta, EU-SILC 2004-2011 (ESeC-88)

		ESeC-88 with old code				
		1	2	4	5	Total
ESeC-88 with new code	1	33,331	0	0	0	33,331
	2	935	8,649	0	0	9,584
	4	5,885	0	110,201	0	116,086
	5	646	0	0	42,072	42,718
	Total	40,797	8,649	110,201	42,072	201,719

Note: ESeC: 1 "Large employers, higher professionals"; 2 "Lower managers/professionals, higher supervisory/technicians"; 3 "Intermediate occupations"; 4 "Small employers and self-employed (non-agriculture)"; 5 "Small employers and self-employed (agriculture)"

Table 22. Number of self-employed respondents that change social class (ESeC) in the new code, persons at active age and in work only, all countries except Malta, EU-SILC 2011-2017 (ESeC-08)

		ESeC-08 with old code					
		1	2	3	4	5	Total
ESeC-08 with new code	1	31,732	0	0	0	0	31,732
	2	2,172	22,407	0	0	0	24,579
	3	0	0	346	0	0	346
	4	3,548	0	0	92,201	0	95,749
	5	814	0	0	0	46,218	47,032
	Total	38,266	22,407	346	92,201	46,218	199,438

Note: ESeC: 1 "Large employers, higher managers/professionals"; 2 "Lower managers/professionals, higher supervisory/technicians"; 3 "Intermediate occupations"; 4 "Small employers and self-employed (non-agriculture)"; 5 "Small employers and self-employed (agriculture)"

Because of this change, the share of the self-employed in the upper (ESeC 1 and 2) and middle class (ESeC 2-6) changes as well. As we have seen before, there is quite some variation in the share of self-employed in the upper class, in 2017 varying from about 20 per cent in Greece and Italy, over around 15 per cent in the Czech Republic, Spain, the Netherlands and Slovakia, down to around 5 per cent in Denmark, Croatia, Malta and Norway. The graph below illustrates the strong variation across countries in the impact of the change in classifying the self-employed on their share in the upper class. While in some countries the change has no effect at all (CH, FI, IE, IS, NL, RS, UK), and in most the impact is rather modest (especially in the case of ESeC-08), it has a very strong effect on the share of the self-employed in Romania, Slovakia, Denmark and Iceland (at least in some years). From a substantive point of view, it would be worthwhile to dig deeper into the question of who the self-employed, without employees, but working in a big firm really are. In any case, for some countries treating them as self-employed without employees rather than large employers, seems to make a non-negligible difference. In other respects, the shift in coding does not seem to make an important difference for most countries, and most observations about the quality of ESeC in EU-SILC remain the same: The average (and often high) ratio of the share of the self-employed in

the ESeC-08 years and in the ESeC-88 years remains roughly the same across countries. Further, only for Italy there is a small increase in the proportion of respondents that is classified differently under ESeC-08 as compared to ESeC-88 in 2011[7]. Finally, also within-country longitudinal variations in the share of self-employed in the middle and upper classes are very similar to those observed with the GESIS code.

Figure 7. Ratio of the weighted share of the self-employed in the upper two ESeC categories with the new coding as compared to the GESIS code, EU-SILC 2007 (ESeC-88) and EU-SILC 2017 (ESeC-08)
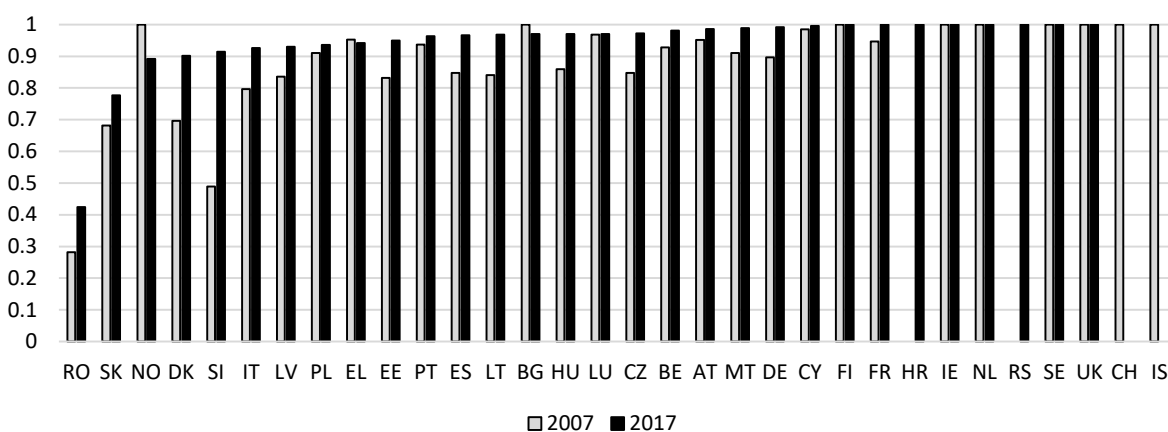
Note: weighted share, people at working age and active only. Countries ordered by 2017 values.

Table 23. Ratio of the weighted share of the self-employed in ESeC categories 3-6 with the new coding as compared to the GESIS code, EU-SILC 2007 (ESeC-88) and EU-SILC 2017 (ESeC-08)
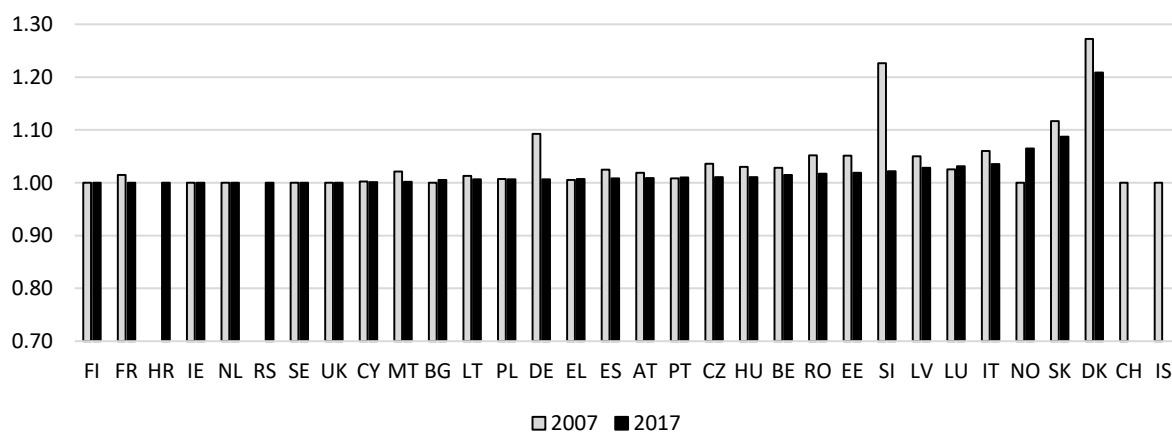
Note: weighted share, people at working age and active only. Countries ordered by 2017 values.

---

[7] The proportion of observations that changes at least one class up-or downwards with the new coding is 0.175 as compared to 0.165 with the GESIS code, in a three-class schema. See Table 9 for more details.

Table 24. The weighted share of the self-employed in the upper two ESeC categories and ESeC categories 3-6, comparing the new code with the GESIS code, EU-SILC 2007 (ESeC-88) and EU-SILC 2017 (ESeC-08)

	ESeC classes 1 and 2				ESeC classes 3-6			
	2007 (ESeC-88)		2017 (eSeC-08)		2007 (ESeC-88)		2017 (eSeC-08)	
	new code	GESIS code	new code	GESIS code	new code	GESIS code	new code	GESIS code
AT	0.04	0.04	0.11	0.11	0.10	0.10	0.22	0.22
BE	0.08	0.09	0.11	0.11	0.24	0.23	0.25	0.25
BG	0.05	0.05	0.08	0.09	0.35	0.35	0.48	0.47
CH	0.10	0.10	.	.	0.23	0.23	.	.
CY	0.06	0.06	0.09	0.09	0.31	0.31	0.33	0.33
CZ	0.12	0.14	0.15	0.16	0.39	0.38	0.46	0.45
DE	0.11	0.12	0.08	0.08	0.13	0.12	0.08	0.08
DK	0.04	0.06	0.05	0.06	0.11	0.09	0.10	0.08
EE	0.05	0.06	0.08	0.08	0.26	0.25	0.39	0.38
EL	0.21	0.22	0.20	0.21	0.61	0.61	0.65	0.65
ES	0.10	0.12	0.14	0.14	0.39	0.38	0.46	0.46
FI	0.08	0.08	0.09	0.09	0.38	0.38	0.44	0.44
FR	0.06	0.07	0.08	0.08	0.21	0.21	0.25	0.25
HR	.	.	0.05	0.05	.	.	0.39	0.39
HU	0.09	0.10	0.11	0.11	0.36	0.34	0.39	0.38
IE	0.08	0.08	0.11	0.11	0.43	0.43	0.34	0.34
IS	0.08	0.08	.	.	0.28	0.28	.	.
IT	0.19	0.24	0.21	0.23	0.44	0.41	0.46	0.45
LT	0.02	0.03	0.06	0.06	0.36	0.36	0.52	0.51
LU	0.06	0.06	0.06	0.06	0.10	0.10	0.15	0.14
LV	0.06	0.08	0.09	0.10	0.24	0.23	0.45	0.44
MT	0.08	0.09	0.04	0.04	0.34	0.33	0.30	0.30
NL	0.10	0.10	0.14	0.14	0.23	0.23	0.29	0.29
NO	0.06	0.06	0.04	0.05	0.21	0.21	0.24	0.23
PL	0.07	0.08	0.09	0.10	0.48	0.48	0.61	0.61
PT	0.14	0.15	0.09	0.09	0.47	0.47	0.38	0.37
RO	0.05	0.17	0.04	0.10	0.65	0.62	0.80	0.78
RS	.	.	0.08	0.08	.	.	0.49	0.49
SE	0.05	0.05	0.08	0.08	0.11	0.11	0.37	0.37
SI	0.03	0.06	0.06	0.07	0.17	0.14	0.24	0.23
SK	0.07	0.10	0.14	0.18	0.27	0.24	0.48	0.44
UK	0.06	0.06	0.09	0.09	0.28	0.28	0.44	0.44

Note: Cells shaded in grey indicate a difference of at least to two points (excluding rounding errors).

6   Concluding remarks

In this note, I briefly assessed some aspects of the data availability and data quality for the individual variables that are required to construct the ESeC classification in EU-SILC, as well as the desired outcome variable. Overall, there seem to be important challenges in terms of data quality and response patterns for comparisons over time, in many countries, as well as comparisons across countries at a single point in time. This can be seen from variations in (clearly non-random) non-response patterns and changes in the share of self-employed among the middle and upper class. These observations should function as a checklist when evaluating observed patterns in within and between income inequality across social classes. When interesting patterns/changes or variations across time occur, it seems wise to check first whether this observation might not be (strongly) confounded by the data limitations mentioned in this note.

In what follows, I discuss somewhat more detailed take-home messages:

Sample restrictions:

- If possible, restrict sample to working age and those who agree they are currently working: this is the sample for which ESeC is in principle defined. It seems best to me to apply these restrictions deliberately. PL040 allows also deriving previous ESeC for those currently unemployed, but only for a (small) subsample. In some countries, ESeC is only defined for people that are working, but not in all. Making the deliberate choice to restrict the target population to those who define themselves as currently an employee or self-employed (PL031), would increase consistency. At the same time, though, we should be well aware that the patterns we will observe might be partially influenced by the changes in the number of unemployed and (early) retired, which probably will not be randomly distributed across social class. Obviously, for some research questions this restriction is not possible, and then these observations should be taken into account when interpreting the findings.
- The data quality of EU-SILC 2004 seems sufficient (or at least not worse than 2005) for most countries that have the data available. Notable exceptions are Finland, which has (much) lower response rates for some of the key variables to construct ESeC-88, resulting in middle and upper classes that consist exclusively of self-employed. Also in Belgium and Estonia some variables (PL150 in BE, PL040 and PL150 in EE) have a substantially lower response rate for 2004, but this does not seem to affect the overall availability and composition of the ESeC variable in both countries. Similarly, in the case of Denmark the years 2012-2015 (incl.) are difficult to use, given that in these years self-employed are completely missing from the middle and upper class. However, employees could still be subject of study. In contrast, for Sweden, 2004 might offer a better point of reference than 2005, given the better data availability and % self-employed in the middle class in that year, which is more in line with the following years than the 2005 situation.
- As far as the selected respondent countries are concerned, our main worry should not be the random selection of individuals (unless one wants to aggregate to the household level), but the non-response among the selected respondents. Average earnings and inequality among the non-respondents may help to understand the patterns we observe. Similar non-response analyses can be carried out for other research questions, and may help to avoid interpreting data artefacts as substantive findings.

For the analysis:

- If one is interested in longitudinal analysis, all countries exhibit a break in series (mostly in 2011, some in 2010) due to the transition of ISCO-88 to ISCO-08. For a longitudinal analysis that includes many countries, it seems best to have a break in series in 2011. Interestingly, for most countries both ISCO-88 and ISCO-08 are available for that year, allowing to generate two ESeC variables for an overlapping year.
- As is the case for the selected respondent countries, non-response patterns should help guide the interpretation of the overall findings.
- Given that the composition of the upper class and middle class in terms of share of employees vs. self-employed is to quite some extent influenced by response-patterns for variables PL150, respectively, PL130, it seems advisable to repeat the analysis for employees separately. One can consider repeating the analysis also for the self-employed, but the sample sizes may then be too small. When doing the analysis for employees-only, one should be very cautious with drawing more general conclusions about social class (including the self-employed) given that the share of the self-employed in the middle and upper social class varies so strongly cross-nationally, and sometimes account for a very sizeable part (of up to more than half) of the middle class.

Alternatives for defining social class:

- If current employment (self-defined current economic status, PL031) is the point of reference for defining the target sample/population, one could wonder why this is not also used for defining social class (PL031 rather than PL040), given that in some cases the classification of self-employed vs. employee is different (about 5% of self-employed in PL040 are considered employee according to PL031). This would be more in line with the computations done in, but deviate from GESIS practice (which in effect only uses PL031 when PL040 is not available). Unfortunately, this practice is only possible from 2009 onwards.
- In addition, one may wonder why the information in PL040 on self-employed with employees and self-employed without employees is not used to make a distinction between these categories, rather than relying solely on the size of the local economic unit (PL130). To me, it would seem more logical to assess the size of the local unit only in the case of self-employed with employees (PL040 == 1), and assume that the size of the local unit is 1 in the case of self-employed without employees. The (non-negligible) group of self-employed without employees, but in a local unit consisting of more than one person, is a very ambiguous category. An application of this alternative approach demonstrates that at least in several countries (most notably Romania) this change in coding would have a non-negligible effect on the classification of the self-employed in general and the share of the self-employed in the upper (and to some extent the middle) class in particular. I am convinced that this is a strategy worth pursuing in future research.

7 Appendix: Remarks by country

The table below summarizes the data issues by country. Crosses in brackets are warnings, crosses without brackets indicate a potentially more severe data issue in the period concerned, while cells shaded in grey indicate 'official' breaks in series (in the at-risk-of-poverty indicator).

Table 25. Overview by country

selecte d resp.	country	2005-2011	2011-2017	remarks
	AT	(x)		PL130 available for very few self-employed in 2007 & 2008 => drop in n for those at work
	BE			Change in weighting in 2012; 2005 SILC armed forces (code 01) have been coded by mistake 11
	BG	x		available from 2007 onwards; break in 2016; PL040 and PL050 contain relatively few observations in 2007, and much lower share of self-employed in 2007; ISCO-88 missing for 2011; PL130 much higher response rate in 2007 than thereafter
	CH			2017 not yet available
	CY			
	CZ			
	DE	x	x	Quota sample until 2008; Strong drop in response rate PL040 in 2007 and strong increase again in 2010. Strong increase/change in response rate PL130 and PL150 over the course of the first period; Since 2015 PL051 only 9 categories.
S	DK	(x)	x	PL130 available at more aggregated level in 2014 (misclassification self-employed in units of 10 and more); Careful with sudden drops / increases in data availability for ESeC88 / ESeC08 in 2006 and 2014/2015. PL130 not available for self-employed in period 2012-2015 (inclusive), so no self-employed with observed ESeC in this period.
	EE			
	EL			
	ES			
S	FI	x		2004: middle and upper class consist of self-employed only; 2009: strong jump in response rates PL040 and PL050; ISCO-88 missing for 2011
	FR	x	(x)	Strong variation in people at work for which social class is available in the first period, in 2011 lower response rate for esec-08
	HR	x		available from 2010 onwards
	HU	(x)		Some undefined minor (?) problems with PL050 in 2005; much lower response rate ESeC-88 in 2005 (among active pop); sudden much higher share of self-employed in 2010 and 2011 and drop in 2014); Jump in response rate for PL130 in 2016
	IE	x	x	2006 strong drop in response rate PL040; ISCO-88 missing for 2011; second period: unclear aggregation of PL051; 2014 jump in response rate of PL130
S	IS		x	Since 2014 PL051 not available; 2017 not yet available.
	IT			

selected resp.	country	2005-2011	2011-2017	remarks
	LT			Careful with changes in share of self-employed.
	LU		▓	Break in 2016; includes strong drop in response rate PL050 and strong increase in response rate PL130
	LV			2009: strong drop in response rate PL130
	MT	x	x	Available from 2007 onwards. PL050 since 2008 available only at 1 digit level
S	NL	x	▓	Low response rate for ESeC-88 in 2005; Break in 2016; 2013-2017 small number allocated to 'weird' PL051 categories. Response rate PL040 jumps in 2006 and 2009. PL130 contains only 5 categories, misclassifying economic units of 10.
S	NO			Sudden temporary drop in response rate PL040 and PL050 in 2014; relatively low ESeC-08 response rate in 2011; careful with changes in ESeC availability and non-random non-response
	PL			
	PT			
	RO			Available from 2007 onwards; 2009: drop in response rate PL130; 2014 sudden temporary drop in response rate for PL051; PL050 not available in 2010, but PL051 (with reduced number of categories) is. In analysis assumed that for 2010 PL050 = PL051.
	RS	x		Available from 2013 onwards. Strong increase in response rate PL130 in 2014
S	SE	▓		Break in series in 2008. Strong drop in response rate PL130 since 2006, and increased strongly again only since 2012. High share of self-employed in 2005; Low share of self-employed in 2011 (ESeC-08). Careful with fluctuations in ESeC availability.
S	SI	x	x	2006: jump in response rate PL130; 2009: strong change response rate PL040 and PL050; Since 2014 PL051 only available at 1 digit level
	SK		x	until 2014 significantly reduced number of categories for PL051
	UK	(x)	▓	2005: unusually high share of self-employed in middle/upper class. Break in 2017; some variation in availability of data in first period (esp. 2009); strong drop in response rate PL150 in 2012.

Note: Official breaks in series are shaded in grey.

References

Rose, D. and Harrison, E. (2007), 'The European socio-economic classification: a new social class schema for comparative European research', *European Societies*, 9: 3, 459-490.