SSOAR

Open Access Repository

www.ssoar.info

Anonymized EU-LFS Microdata for Research: Background, Resources, and Introduction to Data Handling

Mack, Alexander; Lengerer, Andrea; Dickhaut, Theresa

Veröffentlichungsversion / Published Version
Arbeitspapier / working paper

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der

Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-49120-1

gesis

Leibniz-Institut
für Sozialwissenschaften

Anonymized EU–LFS Microdata for Research

Background, Resources, and Introduction to Data Handling

Alexander Mack, Andrea Lengerer &
Theresa Dickhaut

Anonymized EU-LFS Microdata for Research

Background, Resources, and Introduction to Data Handling

Alexander Mack, Andrea Lengerer & Theresa Dickhaut

Inhalt

Summary

This report is intended to aid researchers working with the European Labour Force Survey, which Eurostat disseminates in anonymized form to accredited researchers. The report seeks to fulfil three objectives: Firstly, to provide an overview of the scope of the survey and it's potential for scientific inquiry. Secondly, to review the wide array of resources made available by Eurostat and the services provided by GESIS and provide links for further reading. And finally, to instruct on the use of the microdata by outlining the data structure, the different identifiers and how these can be used for analyses at the household level and to study individuals over multiple quarters. To this aim the report includes a number of illustrated examples alongside Stata code. These examples are mainly intended for researchers with little experience with complex data files, they can also be useful to experienced researchers as they highlight some of the key features of the EU-LFS microdata.

1   European Labour Force Survey

1.1   Overview

The European Labour Force Survey (EU-LFS) is the European Union's prime data source on the labour market and is used to calculate monthly unemployment statistics. With roughly 1.8 million interviews per quarter, it is by far the largest household survey in Europe. The first European Labour Force Survey was carried out in 1960 by the six original member states. To date, 33 countries participate in the EU-LFS; this includes the 28 EU member states, the EFTA countries, Switzerland, Norway and Iceland as well as the candidate countries Macedonia and Turkey (Eurostat 2016a). Microdata for research purposes are available from 1983 onwards.

The EU-LFS is devised as an output harmonized survey. The scope and content of the survey, as well as criteria to ensure data quality and comparability are defined by EU legislation. The national statistical institutes are responsible for the data collection in their respective countries. Eurostat is responsible for integrating the national data files, ensuring data quality and disseminating results, as well as integrated microdata files to researchers. The national statistical institutes are obliged to transmit data and standardized quality reports to Eurostat in a timely manner (Eurostat 2016b). This leaves the participating countries with considerable leeway regarding the method of data collection. While most countries rely on interviews for data collection, others employ national registers as an important data source which is complemented by a survey.

As mentioned above, the topical focus of the EU-LFS is the labour market. It collects detailed information on labour status, employment characteristics, hours worked, second job, previous work experience, search for employment, methods used to find work, education, training and income from the main job. Furthermore, information on labour status and employment for the previous year is collected which means labour market transitions can be studied with the LFS. Additionally, the survey collects detailed information on demographics and the microdata also contain a number of derived variables which describe household characteristics. Over the years the scope of the EU-LFS has steadily increased, and new variables have been added on several occasions. From 1999 onwards the EU-LFS also includes ad hoc modules with specific and recurring topical focuses such as lifelong learning, reconciliation of work and family or transitions into retirement (Eurostat 2014a).

The EU-LFS is used to monitor Europe 2020 targets on employment (European Commision 2015). Additionally, Eurostat disseminates indicators derived from EU-LFS data in the LFS database.[1] The EU-LFS microdata are also made available to accredited researchers free of cost. Due to the large sample sizes, the broad European coverage and the fact that microdata are available from 1983 onward the EU-LFS has become a popular data source for comparative research, with over 100 research proposals submitted to Eurostat in 2014 alone (Bujnowska & Schweikle-Hilgner 2015). Researchers, who are granted access, receive data for all rounds of the LFS from 1983 onwards. Data are delivered as comma separated values (.csv) and include yearly, quarterly and ad hoc files. The yearly files include all variables while the quarterly files only include those variables which are collected in all quarters as defined in regulations 430/2005 and 377/2008 (Eurostat 2014b, p. 62).

---

[1]   http://ec.europa.eu/eurostat/web/lfs/data/database

1.2   Legal Framework

The EU-LFS is carried out on the basis of EU-legislation which gives clear instructions on the information which is to be collected within the EU-LFS, the implementation of the survey, quality standards of the data and how tasks are divided between National Statistical Institutes and Eurostat. These strict and detailed legal boundaries are intended to assure the comparability of transnational results.

There are several general regulations specifying the survey frame and the legal basis of the EU-LFS. Council Regulation (EC) No 577/98 constitutes the principal legal foundation of the EU-LFS which specifies its implementation rules. The regulation determines that the EU-LFS must be implemented as a continuous survey, and provide quarterly and annual results. Regulation (EC) No 1991/2002 of the European Parliament and of the Council declares the continuous adaptation of the EU-LFS to be mandatory from the year 2003 onwards. Regulation (EC) No 2257/2003 establishes the distinction between quarterly and structural variables. In the latter case, the variables are allowed to be surveyed as annual instead of quarterly averages. The Regulation (EC) No 1372/2007 of the European Parliament and of the Council of 23 October 2007 requires that providing information on income is mandatory for all participants.

Regulation (EC) No 1897/2000 details the contents of the survey. It specifies the definition of unemployment and details principles for the formulation of the questions regarding the labour status. The organization of the ad hoc modules is defined by Regulation (EC) No 596/2009. In addition, Regulation (EU) No 545/2014 determines financing provisions and sample conditions for the ad hoc modules.

The scientific access to the EU-LFS data is also regulated by law. Commission Regulation (EC) No 1000/2007 regulates the access to data for researchers from different institutions and organizations. Commission Regulation (EC) No 557/2013 provides information about the form of data and access conditions.

1.3   Implementation

The EU-LFS is devised as an output harmonized survey. A catalogue of variables and a number of quality criteria are defined in EU legislation. The National Statistical Institutes, which are responsible for data collection in their respective countries, are free to collect the data as they see fit as long as the precision and quality criteria are met. The microdata, as well as a standardized quality report, are transmitted to Eurostat. Eurostat integrates the national data and calculates indicators based on these microdata. These indicators are then published in their statistics database. Additionally, an anonymized version of the microdata is prepared by Eurostat and made available to accredited researchers.

Each country has considerable leeway in how they choose to design and conduct their national labour force survey and in many countries, the data collection for the EU-LFS is combined with national data collection efforts. The mode of data collection varies considerably between countries. Some countries rely entirely on telephone interviews while others conduct only face to face interviews and yet others employ mixed mode surveys. In the Nordic countries (Denmark, Finland, Iceland, Sweden and Norway) both survey and register data are utilized for the EU-LFS. There is also considerable variation in questionnaires, item wording and even response categories (see Section 2.3 for details).

In order to better identify short-term trends in the labour market and to better estimate yearly unemployment rates, the EU-LFS is designed as a continuous quarterly survey. The objective is to measure characteristics of the population in an average week in each quarter. This is achieved by uniformly spreading the sample for each quarter over all weeks of a quarter (Eurostat 2015, p. 3). Initially, the labour force survey was conducted in spring (i.e. the 2nd quarter of the year), but based on Regulation 577/98 countries began to transition to a continuous survey during the 1990s (Eurostat 2016b). Since

2014 all countries have switched to a continuous survey. The transition to a continuous survey also has implications for the data collection and structure as all variables must be collected on a yearly basis but only some information needs to be collected in every quarter. Variables collected on a yearly basis are called structural variables.

In most countries, the EU-LFS is implemented as a rotating panel survey, in which individuals are surveyed in multiple quarters (Eurostat 2014c, p. 7). There is considerable variation in the number of panel waves between countries. A commonality among all these designs is that interviews tend to be carried out in subsequent quarters at least once each year. Table 3 in the Appendix provides an overview over the rotation schemes employed throughout Europe for the 2013 EU-LFS.

2   Anonymized European Labour Force Survey

2.1   Coverage

The EU-LFS currently covers 33 countries, including the EU28 countries, the three EFTA countries Iceland, Norway and Switzerland as well as the two candidate countries Former Yugoslav Republic of Macedonia and Turkey (Eurostat 2016a). The oldest available EU-LFS microdata date back to 1983. From 2002 onwards, full coverage of EU28 has been achieved. However anonymized microdata are not available for all countries. Table 1 below provides an overview on availability of the different data files for all participating countries.

Table 1: Availability of yearly, quarterly, and ad hoc microdata files

	Yearly	Quarterly	ad hoc
AT	1995	1995	2002
BE	1983	1983	2002
BG	2000	2000	2003
CH	1996	1996	2003
CY	1999	1999	2002
CZ	1997	1997	2002
DE	2002	2002	2002
DK	1983	1983	1999 (2000 & 2001 n.a.)
EE	1997	1997	2002
GR	1983	1983	1999 (2000 & 2001 n.a.)
ES	1986	1986	2002 (2011 n.a.)
FI	1995	1995	1999 (2000, 2001& 2008 n.a.)
FR	1983	1983	2003
HU	1996	1996	1999 (2000-01 n.a.)
HR	2006	2006	2010
IE	1983	1983	1999 (2000-01 n.a.)
IS	1995	1995	2004 (2008 n.a.)
IT	1983	1983	1999
LV	1998	1998	2003
LT	1998	1998	2002
LU	1983	1983	1999 (2000-02 n.a.)
MT	2009	2009	2009
NL	1983 (1984 & 1986 n.a.)	1983 (1984 & 1986 n.a.)	2002
NO	1995	1995	2001
PL	1997	1997	2002 (2001 n.a.)
PT	1986	1986	1999 (2000-01 & 2003-04 n.a.)
RO	1997	1997	2002
SE	1995	1995	1999
SI	1996	1996	2002
SK	1998	1998	2002
UK	1983	1983	2002

Initially the EU-LFS was conducted as a quarterly survey. It has transitioned to a quarterly continuous survey gradually from 1998 onwards. Between 1983 and 1997, data collection was conducted in spring during the first or second quarter, depending on the respective country. Since 2005 disseminated results represent annual averages of quarterly data for most countries since they are available due to the compulsory quarterly continuous data collection.

2.2   Anonymization Criteria

In order to enable the access to the EU-LFS microdata to the scientific community, several anonymization and aggregation criteria were agreed on between the National Statistical Institutes, the owners of the data, and Eurostat. Anonymization efforts regard several concepts concerning the anonymity of persons and households through deleting or aggregating certain variables. The anonymization criteria are applied to both core as well as ad hoc module datasets. Consequently, the household numbers are randomized per dataset to minimize the risk of disclosure of the statistical units across waves. There are further anonymization and aggregation criteria that are specific to individual member states and ad hoc modules.

The anonymization of certain variables imposes considerable constraints on some research questions. For example, the anonymization criteria applied to age, which is reported in five year age bands, imposes limitations to research on family formation processes or inquiries on the intersections of parenthood and employment. Similarly, the broad clustering of regions for country of birth might be problematic for research on migration. Scholars interested in conducting analysis on the basis of EU-LFS microdata should carefully examine relevant variables to ensure that data are sufficiently precise to meet their needs.

2.3   Comparability Issues

The EU-LFS microdata are delivered as a fully harmonized data file. However, researchers should not make the mistake to take for granted that microdata are fully comparable. Comparability issues can result from changes in methods, questionnaires or sampling designs between waves, or from the differences in implementation of the survey between countries.

Researcher need to be aware of the fact that harmonization in the EU-LFS is achieved via output harmonization, and that such an approach contains a number of pitfalls.[2] Particularly in regards to comparability between countries the different methods and tools employed by National Statistical Institutes throughout Europe can be problematic. Researchers are urged to make use of the available documentation to examine comparability issues.

In regards to design, one major difference between countries is the mode of data collection. Most countries rely exclusively on data from household surveys, a number of countries (mainly the Nordic countries) also utilize registers during data collection. There is also considerable heterogeneity between countries in regards to interview mode. Furthermore, in some countries participation is compulsory, while in others it is voluntary. This leads to large differences in response rates between countries (i.e. 31% in Luxemburg versus 97% in Germany).

---

[2]   For a critical discussion of output harmonization see for example Körner and Meyer (2005) or Hoffmeyer-Zlotnik and Warner (2014).

A further issue which can come up in an output harmonized survey is that of non-matching response categories. Research on comparability of supervisory status in the EU-LFS (Pollack et. al. 2009) reveals considerable differences in the prevalence of supervisory status in Europe, which are most likely attributable to differences in item wording between countries.

Further issues which can affect comparability are breaks in methodology over time. A very prominent example for the LFS is the change from a quarterly to a continuous survey from 1998 onwards. Other potential problems include changes in classifications (e.g. in the NACE classification or from ISCO88 to ISCO08), changes in the codification of variables (e.g. nationality 2004) or changes in sample design. The *Statistics Explained* section on the EU-LFS (Eurostat 2016b) provides information on breaks in methodology.

2.4   Microdata Access

Eurostat provides access to anonymized microdata only for research purposes based on Commission Regulation No 557/2013. Microdata are currently provided to researchers free of charge. The process of requesting access to microdata entails two separate steps. In the first step, a researcher must have his or her institution verified as a research entity. This step is only required once per institution. In a second step a research proposal must be submitted to Eurostat which provides information on individuals involved in the project, its purpose, why microdata is required for the project and how data security will be ensured.[3] After proposals are approved research are provided with DVDs which contain the LFS microdata as comma separated values. The data also come with SAS macros which can be used to read data into SAS. The German Microdata Lab at GESIS also makes so-called Setup files for SPSS and Stata available (see section 2.6). Approval of microdata access tends to be tied to a specific project and time span and researchers are obliged to destroy or delete the data once a project is completed.

Eurostat releases one LFS DVD per year which is named according to the year it was released and is usually sent to researchers at the end of a year or the beginning of the subsequent year. The release for year Y will contain quarterly and yearly data for the year Y-1 and the ad hoc module data for Y-2. In addition to including microdata for the most recent year, Eurostat also continually updates and corrects older microdata with every release. Thus when referencing research results based on EU-LFS microdata citing the data version employed is necessary.

2.5   Documentation provided by Eurostat

The *EU Labour Force Survey Database User Guide* (Eurostat 2014b) is likely the most important document for researchers who intend to utilize the anonymized EU-LFS microdata. Since it is not only intended for use with the anonymized microdata, but also used internally by Eurostat, it sometimes refers to variables not contained in the anonymized microdata. It includes information on all variables, such as the variable and value labels and filter instructions. Additionally, the report provides detailed instructions on how derived variables were constructed. The report also provides details on the data structure of the anonymized microdata and lays out the anonymization criteria agreed upon by Eurostat and the national statistical institutes. The document termed *EU Labour Force Survey Explanatory Notes* (Eurostat 2011), also referred to as User Guide Annex, provides additional information on the concepts and definitions used in the EU-LFS. Furthermore, it details the purpose, definitions and im-

---

[3]   For detailed information see:
      http://ec.europa.eu/eurostat/documents/203647/771732/How_to_apply_for_microdata_access.pdf/82d98876-75e5-49f3-950a-d56cec15b896

plementation rules for all variables in the core EU-LFS. This document was first released 2008 and is updated whenever there are changes to variables.

The report *Main Characteristics of the National Surveys* (Eurostat 2015), which is updated annually and available from 2004 onward, provides information on the implementation of the EU-LFS in participating countries. It includes details on the history of the national LFS, sampling, weighting and data collection and an overview of major changes in methodology. A second document which discusses the national studies is the *Quality report of the European Union Labour Force Survey* (Eurostat 2014c). For every survey round of the EU-LFS since 2004 Eurostat publishes a quality report which summarizes sampling and data collection procedures in different countries and details where national surveys deviate from EU regulations, or where individual variables are unavailable.

Another important resource, which provides interested researchers with information on national surveys, are the national Labour Force Survey questionnaires which are made available by Eurostat.[4] Generally, questionnaires are published in the national language but English language translations are available for almost all countries, but not for every year. These questionnaires are a valuable resource to those interested in the comparability of EU-LFS microdata between countries and over time (see also Section 2.3).

Statistics Explained is an electronic publishing platform created in 2009 by Eurostat.[5] Its main purpose is to provide information about Eurostat's statistics. It can be used as an encyclopedia on European Union statistics, an information portal or as a statistical glossary. Statistics explained hosts a collection of articles on the EU-LFS (Eurostat 2016b). These articles are regularly updated and provide an overview of the EU-LFS. These pages include detailed and up to date information on methodology, data dissemination, legal basis, ad hoc modules and the history of EU-LFS.

2.6   Services provided by the German Microdata Lab

The German Microdata Lab (GML) at GESIS offers services to users of official statistics. Historically the activities of the GML focused on German official microdata and specifically the German Microcensus. The GML has a long-standing and close cooperation in place with Destatis, and assists in preparing the anonymized microdata for the German Microcensus. Furthermore, the German Microdata Lab advises statistical offices and legislators on matters relating to the de facto anonymization of microdata and conducts methodological and substantive research on the basis of official statistics microdata.

The GML also provides services for users of microdata provided by Eurostat, mainly for the EU-LFS and the EU-SILC. Since 2009, the GML organizes regular scientific conferences for users of European official statistics in cooperation with Eurostat. The German Microdata Lab also organizes training courses for researchers interested in utilizing European Official Statistics microdata in their work. These training courses are carried out in cooperation with Eurostat and include an introduction to the data and available documentation as well as hands-on training with actual microdata.[6] Another important service provided by the GML is the Microdata Information System (MISSY), which provides structured metadata at the study and variable level for a number of Eurostat surveys including the EU-LFS.[7] MISSY is designed to aid researchers interested in utilizing these data sources, both during data explo-

---

[4]   http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey_-_methodology#Core_questionnaires
[5]   http://ec.europa.eu/eurostat/statistics-explained/index.php/Main_Page/de
[6]   Materials and presentations from training courses conducted by the GML as part of the Data Without Boundaries project can be found on the project's website http://www.dwbproject.org/events/training.html
[7]   http://www.gesis.org/en/missy/metadata/EU-LFS/

ration and analysis and provides structured metadata on studies and variables. The EU-microdata mailing list serves as an information channel about services, events, and publications pertinent to European microdata from official statistics.[8]

Another important service for users of European microdata provided by the GML are setup files for the EU-LFS and EU-SILC. The setups for the EU-LFS read the data as delivered by Eurostat (as comma separated values) into SPSS or Stata, add variable and value labels and introduce a consistent coding scheme for missing values, which conform to social science standards. Additionally, the setups recode alphanumeric variables into numeric variables (where appropriate) and assign a release variable, which includes the year and month of the data release. The setups for the EU-LFS files consist of different syntax files. One file which includes the value and variable labels and a command file for either the yearly, quarterly or ad hod data which users must execute. In order to run the files users must follow the following steps:

1. Download the files for your data release for either SPSS or Stata from http://www.gesis.org/en/missy/materials/EU-LFS/setups, you will need the label syntax and either the yearly, quarterly or ad hoc syntax files.
2. Open the setup file and specify the path in which .csv files are stored on your computer, the path in which the labels file is stored and the path to which data should be extracted to and the name of the final file. These definitions are only specified once at the beginning of the setup file.
3. Specify the files which you wish to import. Since the Eurostat DVDs contain separate files for each country and year this list can be very long.
4. Run the setup file.

The adjoining readme files provide more detailed information on how to operate the setups and also detail and provide additional information on specifics of the data and any errors.

---

[8] If you are interested in subscribing to the newsletter send an e-mail with 'subscribe' in the subject heading to EU-microdata-l@lists.gesis.org

3   Data Handling in EU-LFS

3.1   Data Structure

The anonymized EU-LFS microdata which Eurostat provides to researchers include a number of different files: yearly, quarterly, ad hoc and special files (Eurostat 2014b: p.62ff). Separate files are delivered for each country and each year.[9]

1)   Yearly files: These files include all variables of the Core LFS. This includes variables collected yearly (the so-called structural variables) and variables which are collected for each quarter. This includes a large number of derived household variables which will be detailed below. For countries in which structural variables are surveyed on a quarterly basis the sample in the yearly files is equivalent to the quarterly files, in other countries the yearly file is a subsample of quarterly files (see Table 3 in the Appendix for details). The yearly files use the following naming convention: two-letter ISO country code year_y (e.g. es1999_y.csv).

2)   Quarterly files: These files include all quarterly variables (as defined in regulations 430/2005 and 377/2008). For an overview of which variables are included in which data file see Table 4 in the Appendix. From 2005 onwards 4 quarterly files are delivered for all countries, with the exception of Switzerland for which 4 quarterly files are available from 2009. These files use the following naming convention: two-letter ISO country code-year-q-x where x identifies the respective quarter (e.g. uk2012Q2.csv).

3)   Ad hoc files: ad hoc module datasets contain all variables included in the Core LFS as well as the ad hoc module variables for a given year.[10] These files are available from 1999 onward and use the same naming convention as the yearly data files: two-letter ISO country code year_y (e.g. es1999_y.csv).

4)   Special files: These files include information on household characteristics for Denmark, Finland, and Sweden. The regular EU-LFS microdata do not contain any information on household characteristics for these countries. However, these files contain only a subsample of the overall EU-LFS survey population in these countries. These files use use the same naming convention as the yearly data files: two-letter ISO country code year_y (e.g. es1999_y.csv).

The EU-LFS is designed as a household survey, in which members of a household are interviewed directly or via proxy. Each observation in the EU-LFS microdata identifies an individual at a specific time point. However, each observation is also part of a household, i.e. individuals are clustered within households. In the quarterly data files individuals are included only once. The yearly data can contain multiple observations of individuals from different quarters however. This is handled differently from country to country, in some countries individuals are only included once in the yearly data files, in other countries yearly files included up to four records per person.

---

[9]   The following matrix provides an overview on the availability of different LFS data files over time
       http://www.gesis.org/en/missy/matrix/datasets/EU-LFS

[10]   See the following link for an overview of topics of the ad hoc modules. http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey_-_ad_hoc_modules

3.2   Identifiers and Household Structure

As explained above, each observation in the EU-LFS microdata corresponds to an individual at a given time point. The EU-LFS contains a number of identifiers which can be used to identify individual observations or households:[11]

HHNUM: the household identifier. This variable is the same across all quarters. If households were surveyed in multiple quarters of a year, you can identify whether a household appears more than once in a yearly data file with this identifier.

QHHNUM: This identifier is identical to HHNUM but adds two digits before the household number in the form of QX where X identifies the quarter to which the observation belongs to.

HHSEQNUM: The sequence number of an individual within the household. This variable can be used in conjunction with either household identifier to identify individuals. It stays the same across waves so that individuals can be followed over quarters.

As identifiers are consistent across the four quarters of a year the anonymized EU-LFS microdata can be used to compare individuals between waves, and can thus be used for longitudinal analysis. However, this applies only to the four quarters of one year (i.e. Q1 and Q4 of 2013). Identifiers of separate yearly files are not consistent, and thus it is not possible to combine data for multiple years (i.e. Q4 of 2013 and Q1 of 2014).

In order to identify a person, one can use the combination of HHNUM and HHSEQNUM. In turn the combination of HHNUM, HHSEQNUM and REFWEEK can identify individual observations (Eurostat 2011: p. 9). In the yearly data files the combination of QHHNUM and HHSEQNUM also uniquely identifies individual observations. However, when using these identifiers multiple observations for one individual for different quarters will be treated as separate cases, which might not always be desirable. As HHNUM and QHHNUM do not uniquely identify cases between countries you will also have to use COUNTRY as an identifier when combining data files for multiple countries to carry out cross-national analyses.

Figure 1 illustrates these identifiers on the basis of two sample households and illustrates some of the specifics of the EU-LFS yearly data files. The sample yearly data file includes two households; the first is from Austria and includes two persons. The household was surveyed in all four quarters of 2012 and thus four separate records for each member are included, one for each wave in which it was interviewed. The second household is from Spain and includes four persons. It shares the same household number with our Austrian household, which is why cross country analyses always have to consider country as an identifier. However, this household is only included in the yearly file for one quarter, as are all other households from Spain. For an overview of how often individual observations can appear in the yearly files see Table 3 in the Appendix.

---

[11]   Note that these variables are not universally available for all rounds of the EU-LFS or for all countries. Thus, much of the information here is not applicable to older EU-LFS microdata, for more detailed information on availability of household information consult the Guidelines (Eurostat 2015c: p.41). When attempting to conduct any analyses which make use of household information for Finland, Denmark or Sweden the special files should be used. Data for Switzerland also does not contain any information on the household.

YEAR	COUNTRY	HHNUM	QHHNUM	HHSEQNUM	REFWEEK
2012	AT	3001	Q13001	1	7
2012	AT	3001	Q13001	2	7
2012	AT	3001	Q23001	1	17
2012	AT	3001	Q23001	2	17
2012	AT	3001	Q33001	1	36
2012	AT	3001	Q33001	2	36
2012	AT	3001	Q43001	1	47
2012	AT	3001	Q43001	2	47
2012	ES	3001	Q23001	1	22
2012	ES	3001	Q23001	2	22
2012	ES	3001	Q23001	3	22
2012	ES	3001	Q23001	4	22

Figure 1: Sample households from yearly data file (fictitious data)

On the basis of these identifiers one can also generate a unique person identifier for the yearly data. However, as the variable QUARTER takes on the value _A for all cases in the yearly data some recoding is required. In a first step one has to examine if any identifiers were assigned twice. The Stata Code below will remove cases which are erroneously coded. In both 2014 and 2013 only two cases can be observed for the entire EU.

Stata Code:

```
duplicates tag year country qhhnum hhseqnum, gen(dup)

tab country dup

drop if dup==1

drop dup
```

In a second step a new QUARTER variable is generated by extracting the second digit of QHHNUM.

```
drop quarter

gen quarter=.

recode quarter .=1 if substr(qhhnum,2,1)=="1"

recode quarter .=2 if substr(qhhnum,2,1)=="2"

recode quarter .=3 if substr(qhhnum,2,1)=="3"

recode quarter .=4 if substr(qhhnum,2,1)=="4"
```

These variables can then be used to generate a unique person identifier. This allows users to identify how often individual cases appear in the yearly data files.

```
egen uniqid = group(country year hhnum hhseqnum)

isid uniqid quarter

duplicates tag uniqid, gen(dup)

tab country dup
```

In addition to the identifier variables the EU-LFS also contains a number of variables which detail the relationship of household members to each other.

HHLINK: Defines the relationship to reference person in the household. Values are assigned as follows:

> 1 Reference person
>
> 2 Spouse (or cohabiting partner) of reference person
>
> 3 Child of reference person (or of his/her spouse or cohabiting partner)
>
> 4 Ascendant relative of reference person (or of his/her spouse or cohabiting partner)
>
> 5 Other relative
>
> 6 Other

HHSPOU: Points to the spouse or cohabiting partner of the person. The spouse is identified via his or her sequence number in the household.

HHFATH: This variable identifies a person's father via the sequence number in the household.

HHMOTH: This variable identifies a person's mother via the sequence number in the household.

YEAR	COUNTRY	QHHNUM	HHSEQNUM	HHLINK	HHSPOU	HHFATH	HHMOTH
2012	ES	Q23001	1	1	2	0	0
2012	ES	Q23001	2	2	1	0	0
2012	ES	Q23001	3	3	0	1	2
2012	ES	Q23001	4	3	0	1	2
2012	ES	Q14001	1	1	2	0	0
2012	ES	Q14001	2	2	1	0	0
2012	ES	Q14001	3	5	0	0	0
2012	ES	Q14001	4	5	0	0	3

Figure 2: Sample households from yearly data file (fictitious data)

Figure 2 illustrates these variables for a few sample households. Household 3001 is included in this example again, and we can see that it is a classic four person nuclear family with father, mother and children. This information can largely be deduced from the HHLINK variable. However, we do not know for certain whether the person with the HHSEQNUM=2 is actually the children's mother. This information can only be deduced from HHMOTH and HHFATH. The second example household in Figure 2 (Q14001) further illustrates why additional variables are required to deduce the relationship between household members. In this household the household reference person (HHSEQNUM=1) and the person with sequence number 2 are a couple. All other persons in the household are other relatives of the reference person, but not direct descendants, according to HHLINK. The fact that person 4 reports that person 3 is the persons mother does provide us with additional information on the relationship of these persons and can be of use if one is interested in identifying families (i.e. parents with children) within households.

The EU-LFS also contains a number of variables which classify households and provide additional information on their characteristics. HHNBWORK, for example reports the number of employed persons of 15 years or older in the household, and HHNBCH2 counts the number of persons under 2 years of age in the household. However, many details cannot be directly deduced from these variables. Particularly variables referring to children in the household do not provide any information about the relationship of individual household members to these children. The only variable which directly addresses one's own children is HHCHILDR. This variable provides information on whether or not a person's own children live in the household. But no variables which provide information about the number or age of own children are included in the EU-LFS microdata. Fortunately this information can be derived on the basis of the HHFATH and HHMOTH variables. As this requires some complex data manipulations the example below outlines how to count the number of own children.

First, one must check whether or not a person has children in the household, and whether these children can be classified as children. This is achieved by generating a set of new variables (mpers) which identifies whether a person is the parent of other persons in the household. Then, based on the mpers variables the number of children for each person in the household is counted. For the example households presented in Figure 2 this would look as follows. Note that for the sake of simplicity this example only examines mothers:

QHHNUM	HHSEQNUM	HHMOTH	kidcount	mpers1	mpers2	mpers3	mpers4
Q23001	1	0	0	.	.	2	2
Q23001	2	0	2	.	.	2	2
Q23001	3	2	0	.	.	2	2
Q23001	4	2	0	.	.	2	2
Q14001	1	0	0	.	.	.	3
Q14001	2	0	0	.	.	.	3
Q14001	3	0	1	.	.	.	3
Q14001	4	3	0	.	.	.	3

Figure 3: Counting number of own children (fictitious data)

This is achieved by programing a loop[12] which repeats a set of commands for all persons in each household (4 in the example, note that when running this command over actual data you will have to tabulate the maximum household size and then define the upper bound accordingly). In a first step the loop generates a set of variables (mtemp) for each person in the household. These variables identify the mother of each individual in the household in numbered variables. In a second step the information contained in mtemp is applied to all cases of the household (see Figure 3). This intermediary step is required only for step three, in which the kidcount variable is increased by one for each person in the household which is the child of a given person.

Stata Code

```
***Counting the number of own children in the household***

***Only counting children not classified as adults in hhpers

tab hhseqnum

gen kidcount=0

foreach i of num 1/4 {

gen mtemp`i'=hhmoth if hhseqnum==`i' & hhpers!=4 //Step1

egen mpers`i'=mean(mtemp`i'), by(year country qhhnum) //Step2

replace kidcount=kidcount+1 if mpers`i'==hhseqnum //Step3

}

***Drop temporary variables and label kidcount***

drop mtemp* mpers*

label var kidcount "Number of own children in HH"

tab kidcount
```

Another possibility for conducting analyses at the level of households or families is to merge members of families into one file. By merging together different members of families or households one can generate family or partner files. The example below will outline how to generate a partner file, and will generate a variable indicating the work arrangements in couples based on the ILOSTAT and FTPT variables.[13] This example utilizes the yearly data file for the 2013 Austrian EU-LFS and limits itself to households surveyed in the second quarter to avoid duplicates. The example will be illustrated on the basis of fictitious sample households. Stata code for replicating the exercise is provided below.

Figure 4 shows a few fictitious partner households. The first household (HHNUM=105) is a 3 person household which consists of man woman and one female child (HHCHILDR=1 for both man and woman). The man is employed full time (ILOSTAT=1 and FTPT=1) while the woman is inactive (ILOSTAT=3) as is the child (ILOSTAT=3). Thus this is an example of a classic male breadwinner household. The second household (HHNUM=118) includes 4 individuals: a couple and two other persons, most likely their

---

[12] For an introduction to looping in Stata see: http://www.ats.ucla.edu/stat/stata/modules/acrossvars.htm

[13] This example was taken from the European introductory workshop to the EU-SILC and the EU-LFS held in Manchester in 2011 (Wirth and Lengerer 2011).

male children. The man is unemployed (ILOSTAT=2) and the partnered woman is working part time (ILOSTAT=1 & FTPT=2). The other two persons in the household are economically inactive. The third household (HHNUM=130) is a two person household. Man and woman are partnered with each other and both are employed full time (ILOSTAT=1 & FTPT=1).

COUNTRY	HHNUM	HHSEQNUM	HHSPOU	SEX	ILOSTAT	FTPT	HHCHILDR
AT	105	1	2	1	1	1	1
AT	105	2	1	2	3	.	1
AT	105	3	0	2	3	.	2
AT	118	1	2	1	2	.	1
AT	118	2	1	2	1	2	1
AT	118	3	0	1	3	.	2
AT	118	4	0	1	3	.	2
AT	130	1	2	1	1	1	2
AT	130	2	1	2	1	1	2

Figure 4: Couple households (fictitious data)

Before joining partners in one row of the data file we must first generate separate files for men and women. In a first step we will generate a file which includes only partnered women (SEX=2 & HHSPOU>0). In order to be able to join these women with their partners two data manipulations are required. First variables needed for our later analysis must be renamed to indicate that they refer to women and are marked with an _W. Secondly the HHSEQNUM is removed and instead HHSPOU is renamed to HHSEQNUM. The reasoning behind doing so is that we later want to merge women to men and thus match HHSPU to HHSEQNUM. Note that this example only works for heterosexual couples. Figure 5 illustrates what this looks like for our example data.

COUNTRY	HHNUM	HHSEQNUM	SEX	ILOSTAT_W	FTPT_W	HHCHILDR
AT	105	1	2	3	.	1
AT	118	1	2	1	2	1
AT	130	1	2	1	1	2

Figure 5: Partnered women (fictitious data)

In the next step a file for partnered men is generated analogous to the one for women. With two differences: ILOSTAT and FTPT receive the suffix _M, and HHSEQNUM is not altered. Additionally, this data set does not include HHCHILDR.

COUNTRY	HHNUM	HHSEQNUM	SEX	ILOSTAT_M	FTPT_M
AT	105	1	1	1	1
AT	118	1	1	2	.
AT	130	1	1	1	1

Figure 6: Partnered men (fictitious data)

These two files are then joined with the help of a 1:1 merge (i.e. each partnered man is assigned a partnered woman). The identifiers COUNTRY, HHNUM and HHSEQNUM are used to merge. Note that in this example COUNTRY isn't actually required. However it is needed when data from multiple countries are included in data files.

COUNTRY	HHNUM	HHSEQNUM	ILOSTAT_M	FTPT_M	ILOSTAT_W	FTPT_W	HHCHILDR	empat
AT	105	1	1	1	3	.	1	1
AT	118	1	2	.	1	2	1	4
AT	130	1	1	1	1	1	2	3

Figure 7: Partner file men (fictitious data)

On the basis of ILOSTAT_M, FTPT_M, ILOSTAT_W, and FTPT_W the variable working arrangements in partnerships (empat) can now be generated easily. Figure 8 displays the results for the second quarter of the 2013 Austrian EU-LFS.
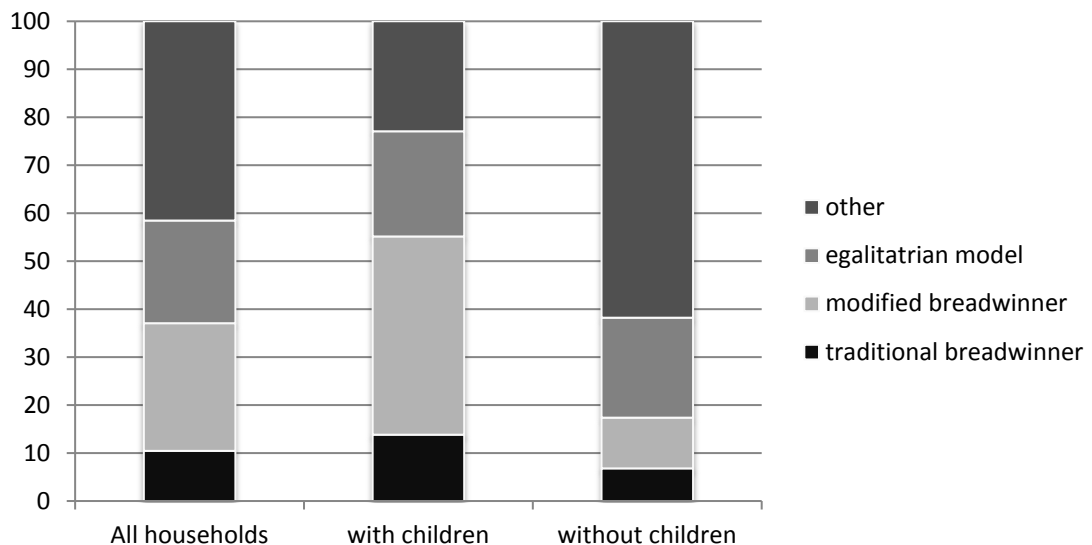
Figure 8: Working arrangements in partnered households (2013 Austrian EU-LFS 2nd quarter)

Stata Code:

```
*Step 1 generate a file which includes only partnered women
cd "C:\yourworkingdirectory\"
use at2013_y.dta, clear
drop quarter
gen quarter=.
recode quarter .=1 if substr(qhhnum,2,1)=="1"
recode quarter .=2 if substr(qhhnum,2,1)=="2"
recode quarter .=3 if substr(qhhnum,2,1)=="3"
recode quarter .=4 if substr(qhhnum,2,1)=="4"
keep if quarter==2 & sex==2 & hhspou>0
keep country hhnum hhspou ilostat ftpt hhchildr
rename (hhspou ilostat ftpt) (hhseqnum ilostat_w ftpt_w)
save partneredwomen.dta, replace


*Step 2: repeat procedure for men
use at2013_y.dta, clear
drop quarter
gen quarter=.
recode quarter .=1 if substr(qhhnum,2,1)=="1"
recode quarter .=2 if substr(qhhnum,2,1)=="2"
recode quarter .=3 if substr(qhhnum,2,1)=="3"
recode quarter .=4 if substr(qhhnum,2,1)=="4"
keep if quarter==2 & sex==1 & hhspou>0


keep country hhnum hhseqnum ilostat ftpt


rename (ilostat ftpt) (ilostat_m ftpt_m)
save partneredmen.dta, replace
```

```
*Step 3: merge men and women

sort country hhnum hhseqnum

merge 1:1 country hhnum hhseqnum using partneredwomen.dta

tab _merge

keep if _merge==3

drop _merge


*Step 4: Working arrangements

gen empat=.

recode empat .=1 if ilostat_m==1 & ftpt_m==1 & ilostat_w==3

recode empat .=2 if ilostat_m==1 & ftpt_m==1 & ilostat_w==1 &
ftpt_w==2

recode empat .=3 if ilostat_m==1 & ftpt_m==1 & ilostat_w==1 &
ftpt_w==1

recode empat .=4 if (ilostat_m==2 | ilostat_m==3) | ftpt_m==2 | ilo-
stat_w==2


label define empat 1 "traditioanl breadwinner model" 2 "modified
breadwinner model" 3 "egalitatrian model" 4 "other"

label val empat empat

tab empat hhchildr, col nof

saving couples.dta, replace
```

3.3 Longitudinal Analysis

The EU-LFS also offers some possibilities for longitudinal analysis. For one the continuous data struc-ture which is in place in all countries of the EU-LFS today allows studying transitions between quarters either on the basis of the quarterly or the yearly data files (see Table 3 in the Appendix for an overview of which countries include multiple observations in the yearly file and which don't). However, if one wishes to study transitions between quarters the COEFF weight is not truly applicable as it is only configured to be representative of the current quarter. Thus the lack of a longitudinal weight is a considerable limitation of the EU-LFS anonymized microdata.[14]

---

[14] Eurostat has been working on longitudinal weights for the EU-LFS over the last few years (Espelage & Kiiver 2015) and has begun to make available flow statistics in its statistics database (Eurostat 2016c).

Additionally the EU-LFS includes retrospective questions which survey the respondent's situation one year prior (WSTAT1Y, STAPRO1Y, NACE1Y1D, NA111y1d, NA111YS CONTRY1Y REGION1Y). With the help of these variables it is possible to examine changes between years on a number of selected variables.

Two examples for studying transitions into unemployment will be presented below. The first example demonstrates how transitions into unemployment between years can be calculated on the basis of the retrospective information provided in WSTAT1Y. The second example examines transitions into unemployment on the basis of quarterly data and demonstrates how LFS microdata can be utilized to study seasonal labour market trends.

The first example utilizes the combination of MAINSTAT and WSTAT1Y which assess respondent's labour status at the time of the interview and one year prior. This example will study those who report being unemployed one year prior to the survey and examine their employment status in the year of the survey. The example compares transitions out of unemployment for men and for women in Austria, Spain and Greece.

Stata Code:

```
cd "C:\yourworkingdirectory\"

use at2013_y.dta, clear

append using lv2013_y.dta

append using es2013_y.dta


gen transout=.

recode transout .=1 if wstat1y==2 & mainstat==1

recode transout .=2 if wstat1y==2 & mainstat==2

recode transout .=3 if wstat1y==2 & mainstat==3

recode transout .=3 if wstat1y==2 & mainstat==4

recode transout .=5 if wstat1y==2 & (mainstat==5 | mainstat==7 | mainstat==8)

label define transout 1 "employed" 2 "unemployed" 3 "student" 4 "retired" 5 "inactive"

label val transout transout

tab country transout

tab country transout [aweight=coeff], row nof

tab country transout [aweight=coeff] if sex==1, row nof

tab country transout [aweight=coeff] if sex==2, row nof
```
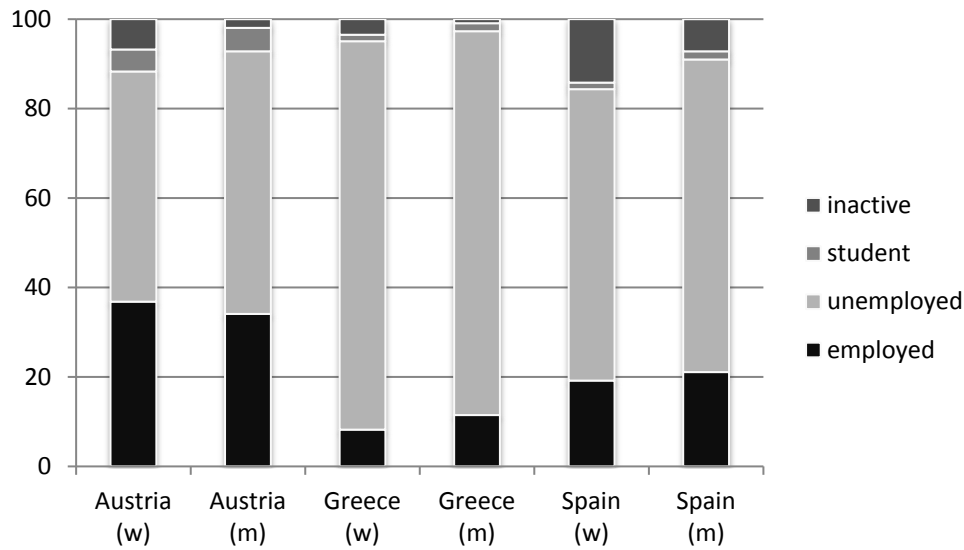
Figure 9: Transition out of unemployment for men and women (2013 EU-LFS)

The second example will examine transitions into employment on the basis of the quarterly EU-LFS files. The example presented below compares transitions into employment on the basis of 2013 EU-LFS data for Austria and Greece. Since MAINSTAT is not available in the quarterly data WSTATOR will be used to generate transitions. The example below is very basic in nature, but the code provided below can easily be adapted to examine other types of transitions as well. In a first step the quarterly files for Austria and Italy are combined into a single data file.

Stata Code:

```
*Step 1: Combining quarterly files
cd "C:\yourworkingdirectory\"
use at2013q1.dta, clear
append using at2013q2.dta
append using at2013q3.dta
append using at2013q4.dta
append using gr2013q1.dta, force
append using gr2013q2.dta, force
append using gr2013q3.dta, force
append using gr2013q4.dta, force
save 2013quarterly.dta
```

In the second step the data is configured as panel data with the help of the tsset command. Once data has been tsset in Stata it becomes much easier to conduct panel analysis due to the lead and lag (t. and l.) operators. However, in order to define a dataset as a panel we need a unique identifier for individuals (uniqid is constructed analogous to section 3.2) as well as a time identifier, in this case quarter.

Stata Code:

```
*Step 2: Using identifiers for tsset
egen uniqid = group(country year hhnum hhseqnum)
isid uniqid quarter
tsset uniqid quarter
```

Before examining transitions into employment, the lag operator is used to examine transitions between quarters. Remember, that due to the different rotation schemes used in the EU-LFS not all persons are interviewed in all four quarters of a year. The code below identifies cases that were interviewed in subsequent quarters. These cases will serve as the sample for the subsequent analysis.

Stata Code:

```
*Step 3: How many adjacent quarters?
gen trans=0
recode trans 0=1 if quarter==2 & l.quarter==1
recode trans 0=1 if quarter==3 & l.quarter==2
recode trans 0=1 if quarter==4 & l.quarter==3

tab trans quarter
table country quarter, c(m trans)
```
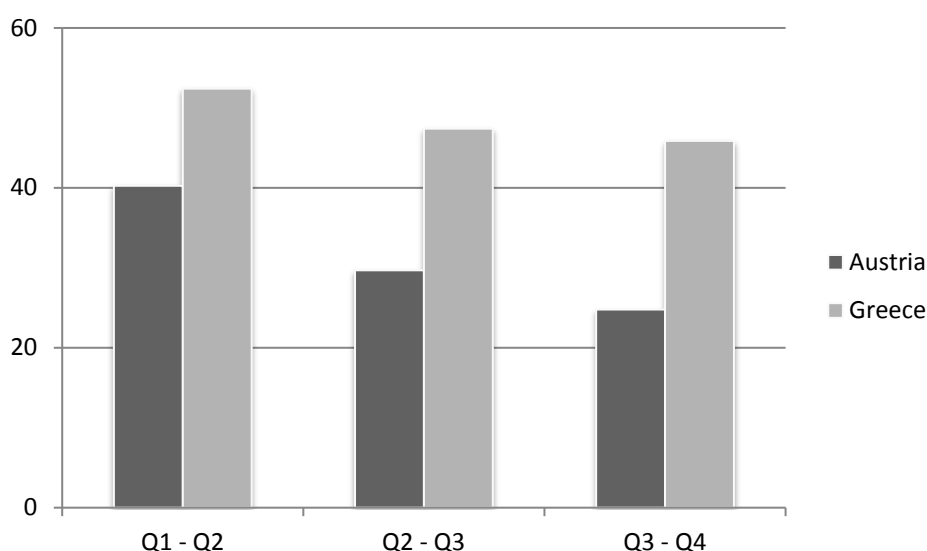

Figure 10: Percentage of cases interviewed in successive quarters (2013 EU-LFS)

Finally, transitions into employment can be calculated. The analysis only considers the working age population, and in accordance with Eurostat's definition of the labour force, does not include individuals in compulsory military service. Labour status during the reference week (WSTATOR) is used to operationalize changes between quarters. The code below makes use of the lead operator (f.) in order

to be able to calculate transitions into employment as proportion of the not working population in the previous quarter. No significant variation between quarters can be observed.

Stata Code:

```
*Step 4: Calculate transitions by quarter and examine differences
between countries


gen inactive=0 if wstator!=4 & age>=17 & age<=67

recode inactive 0=1 if wstator==5


gen transempl=.

recode transempl .=0 if wstator!=. & f.wstator!=. & wstator!=4 &
f.wstator!=4 & age>=17 & age<=67

recode transempl 0=1 if (f.wstator!=5) & wstator==5


tab quarter transempl if inactive==1, row

bysort country: tab quarter transempl if inactive==1, row  nof
```
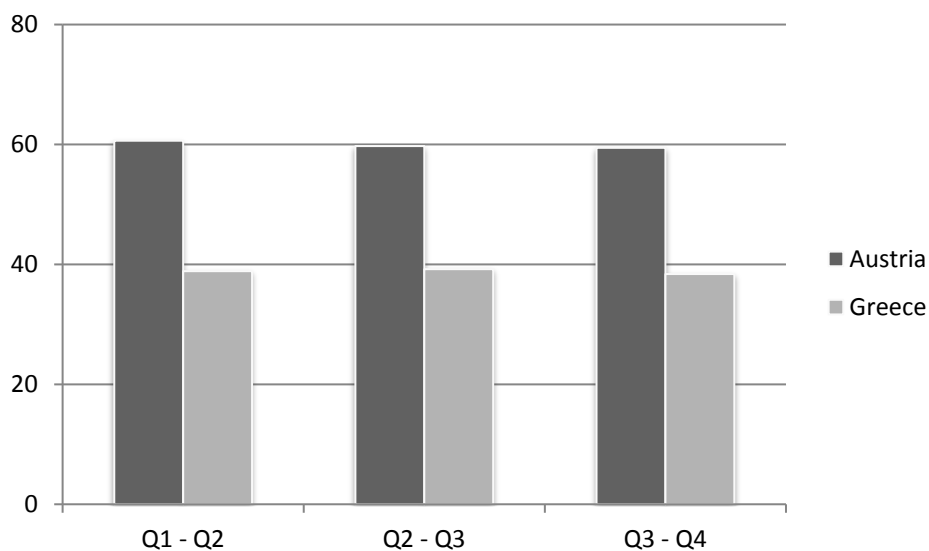

Figure 11: Transitions from inactive to employed between quarters in % (2013 EU-LFS)

4 Sources

Bujnowska, A. &, Schweikle-Hilgner, A. (2015). Access to EU microdata: an overview of the research proposals submitted to Eurostat. Presentation held at the 4th EU User Conference in Mannheim.

Council Regulation (EC) No 577/98 Official Journal of the European Communities L 77/3. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31998R0577&from=EN

Commission Regulation (EC) No 430/2005 of 15 March 2005. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32005R0430&from=EN

Commission Regulation (EC) No 377/2008. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32008R0377&from=EN

Commission Regulation (EU) No 557/2013. Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R0557&from=EN

European Commission (2015). Europe 2020 targets. Retrieved from http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/targets/index_en.htm

Eurostat (2011). EU Labour Force Survey Explanatory Notes (to be applied from 2012Q1 onwards)

Eurostat (2014a). Statistics explained – European Labour Force Survey. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey

Eurostat (2014b). EU Labour Force Survey Database User Guide. Version: December 2014.

Eurostat (2014c). Quality report of the European Union Labour Force Survey 2013. *Eurostat Statistical Working Papers.*

Eurostat (2015). Labour force survey in the EU, candidate and EFTA countries. Main characteristics of national surveys, 2014. *Eurostat Statistical Working Papers.*

Eurostat (2016a). Main features of the EU-LFS and scope of data. [Online] Available at: http://ec.europa.eu/eurostat/web/lfs/overview

Eurostat (2016b). Statistics Explained. EU Labour Force Survey - main features and legal basis. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey_–_main_features_and_legal_basis

Eurostat (2016c). Statistics Explained. Labour market flow statistics in the EU. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Labour_market_flow_statistics_in_the_EU

Espelage, F. & Kiiver, H, (2015). Flow Statistics- Work towards longitudinal analyses based on the LFS. Presentation held at the 4th EU User Conference in Mannheim.

Hoffmeyer-Zlotnik, J. H., & Warner, U. (2014). *Harmonising Demographic and Socio-Economic Variables for Cross-National Comparative Survey Research.* Springer.

Körner, T., & Meyer, J. (2005). Official Statistics: Experiences from the Federal Statistical Office Germany. In J. H. P. Hoffmeyer-Zlotnik & J. A. Harkness (Eds.), *ZUMA-Nachrichten Spezial Band 11, Cross-National Research* (149-162). Mannheim: ZUMA.

Pollak, R., Wirth, H., Weiss, F., Bauer, G., & Müller, W. (2009). On the comparative measurement of supervisory status using the examples of the ESS and the EU-LFS. In B. Pfau-Effinger, S. Sakač Magdalenić, & C. Wolf (Eds.), *International vergleichende Sozialforschung: Ansätze und Messkonzepte unter den Bedingungen der Globalisierung* (pp. 173-206). Wiesbaden: VS: Verlag für Sozialwissenschaften.

Official Journal of the European Communities (2000). Commission Regulation (EC) No 1897/2000 of 7 September 2000 implementing Council Regulation (EC) No 577/98 on the organisation of a labour force sample survey in the Community concerning the operational definition of unemploy-

ment. Retrieved from
http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32000R1897&from=EN

Official Journal of the European Communities (2002). Regulation (EC) No 1991/2002 of the European
Parliament and of the Council of 8 October 2002 amending Council Regulation (EC) No 577/98
on the organisation of a labour force sample survey in the Community. Retrieved from
http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002R1991&from=EN

Official Journal of the European Communities (1998). Council Regulation (EC) No 577/98 of 9 March
1998 on the organisation of a labour force sample survey in the Community. Retrieved from
http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31998R0577&from=EN

Official Journal of the European Union (2003). Regulation (EC) No 2257/2003 European Parliament
and Council of 25 November 2003 amending Council Regulation (EC) No 577/98 on the organi-
sation of a labour force sample survey in the Community to adapt the list of survey characteris-
tics. Retrieved from
http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32003R2257&from=EN

Official Journal of the European Union (2007). Regulation (EC) No 1372/2007 of the European Parlia-
ment and Council of 23 October 2007 amending Council Regulation (EC) No 577/98 on the or-
ganisation of a labour force sample survey in the Community. Retrieved from http://eur-
lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32007R1372&from=EN

Official Journal of the European Union (2009). Regulation (EC) No 596/2009 of the European Parlia-
ment and of the Council of 18 June 2009 adapting a number of instruments subject to the pro-
cedure referred to in Article 251 of the Treaty to Council Decision 1999/468/EC with regard to
the regulatory procedure with scrutiny. Retrieved from
http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32009R0596&from=EN

Official Journal of the European Union (2014). Regulation (EU) No 545/2014 of the European Parlia-
ment and of the Council of 15 May 2014 amending Council Regulation (EC) No 577/98 on the
organisation of a labour force sample survey in the Community. Retrieved from
http://eur-lex.europa.eu/legalcontent/EN/TXT/PDF/?uri=CELEX:32014R0545&from=EN

Wirth, H & Lengerer, A. (2011). Practical session: Exploring EU-LFS. Training course held at the Europe-
an workshop to introduce the EU Survey of Income and Living Conditions and the EU Labour
Force Survey data in Manchester.

5 Appendix

Table 2: EU-LFS 2013 (2014 release), Sample Sizes for different data files

Country	Yearly	Q1	Q2	Q3	Q4	Ad hoc	Special
Austria	180,923	45,212	45,240	45,589	44,882	18,714	–
Belgium	96,519	24,357	23,430	24,102	24,630	23,430	–
Bulgaria	33,240	32,243	30,683	31,630	34,023	33,240	–
Switzerland	71,705	36,544	35,936	35,392	32,022	12,151	–
Cyprus	42,977	10,919	10,950	10,711	10,397	10,950	–
Czech Republic	42,868	56,865	56,088	55,647	55,401	33,331	–
Germany	478,173	111,640	122,617	118,451	125,465	–	–
Denmark	143,312	35,981	35,868	35,709	35,754	35,868	40,303
Estonia	24,315	6,221	6,138	6,036	5,920	12,058	–
Spain	111,056	170,854	171,909	170,601	171,034	105,786	–
Finland	53,392	34,753	34,580	34,905	34,105	33,466	52,503
France	480,641	126,718	122,428	115,033	116,462	26,829	–
Greece	245,180	61,368	60,814	61,332	61,666	60,814	–
Croatia	35,784	9,352	9,144	8,644	8,644	9,144	–
Hungary	259,318	65,571	64,786	64,589	64,372	64,786	–
Ireland	209,237	55,067	53,249	49,518	51,403	53,249	–
Iceland	12,399	3,126	3,103	3,088	3,082	–	–
Italy	611,255	155,047	153,317	150,434	152,457	153,317	–
Lithuania	61,126	15,258	14,971	15,371	15,526	14,971	–
Luxembourg	16,139	3,938	4,132	4,157	3,912	16,139	–
Latvia	39,011	9,451	9,430	9,893	10,237	9,430	–
Malta	25,597	6,318	6,500	6,537	6,242	6,818	–
Netherlands	83,840	113,295	112,093	108,814	108,860	–	–
Norway	23,669	20,524	20,192	19,804	20,160	14,303	–
Poland	391,150	101,330	98,069	96,012	95,739	66,039	–
Portugal	159,659	40,158	39,361	39,647	40,493	39,361	–
Romania	228,549	58,220	57,251	56,232	56,846	56,663	–
Sweden	241,152	61,360	60,443	59,740	59,609	29,891	36,113
Slovenia	61,223	15,031	15,215	15,486	15,491	15,215	–
Slovak Republic	100,810	25,349	25,281	25,091	25,089	25,281	–
United Kingdom	84,041	96,783	95,525	95,559	97,088	84,041	–

Table 3: Rotation and relation of yearly and quarterly files

	Rotation[1]	Sum of Quarterly files=Yearly file[2]	Maximum number of individual observations in yearly data[2]
Belgium	no	yes	4
Bulgaria	2-(2)-2	no	1
Czech Republic	5-	no	1
Denmark	2-(2)-2	yes	4
Germany	4-(annual)	yes	1
Estonia	2-(2)-2	no	2
Ireland	5-	yes	4
Greece	6-	yes	4
Spain	6-	yes	1
France	6-	yes	4
Croatia	2-(2)-2	yes	2
Italy	2-(2)-2	yes	2
Cyprus	6-	yes	4
Latvia	2-(2)-2	yes	2
Lithuania	2-(2)-2	yes	2
Luxembourg	no	yes	1
Hungary	6-	yes	4
Malta	2-(2)-2	yes	2
Netherlands	5-	no	1
Austria	5-	yes	4
Poland	2-(2)-2	yes	2
Portugal	6-	yes	4
Romania	2-(2)-2	yes	2
Slovenia	3-(1)-2	yes	3
Slovak Republic	5-	yes	4
Finland	3-(1)-2	no	1
Sweden	8-	yes	4
United Kingdom	5-	no	1
Iceland	3-(2)-2	yes	3
Norway	8-	no	1
Switzerland	2-(2)-2	no	1

1 Main characteristics of national surveys 2013 p.69; 2 own calculations 2013 EU–LFS (2014 release)

Table 4: EU-LFS 2013 (2014 release), Variables included in different data files

Variable	Variable label	Yearly	Quart. 1-4	Ad hoc	Special
QHHNUM	Serial No. of household in each quarter	X	X	X	X
COEFF	Weighting factor	X	X	X	X
QUARTER	Reference quarter	X	X	X	X
HHNUM	Serial No. of household	X	X	X	X
REC	Reference Code	X	X		
HHSEQNUM	Sequence No. in household	X	X	X	X
REFYEAR	Year of survey	X	X	X	X
REFWEEK	Reference week	X	X	X	X
INTWEEK	Interview week	X	X	X	X
DEGURBA	Degree of urbanization	X	X	X	X
HHTYPE	Type of household (transmitted)	X		X	X
HHINST	Type of institution	X	X	X	X
INTWAVE	Sequence No. of the survey wave	X	X	X	X
INTQUEST	Questionnaire used	X	X	X	X
REM	Reference month	X	X	X	X
YEAR	Fixed reference year	X	X	X	X
HHPRIV	Type of household (derived)	X	X	X	X
HHLINK	Relationship to reference pers. in household	X		X	X
HHSPOU	Sequence No. spouse/cohabiting partner	X		X	X
HHFATH	Sequence No. of father	X		X	X
HHMOTH	Sequence No. of mother	X		X	X
SEX	Sex	X	X	X	X
AGE	Age	X	X	X	X
MARSTAT	Marital status	X		X	X
PROXY	Nature of participation in the survey	X	X	X	X
WSTATOR	Labour status during the reference week	X	X	X	X
NOWKREAS	Reason for not working though having a job	X	X	X	X
STAPRO	Professional status	X	X	X	X
SUPVISOR	Supervisory responsibilities	X		X	X
SIZEFIRM	No. of persons working at the local unit	X		X	X
YSTARTWK	Year start work for employer/self-employed	X	X	X	X
MSTARTWK	Month start work for employer/self-empl.	X	X	X	X
WAYJFOUN	Involvement of public employment office in finding the present job	X		X	X
FTPT	Full-time / Part-time distinction	X	X	X	X
FTPTREAS	Reasons for the part-time work	X		X	X
TEMP	Temporary work: Permanency of the job	X	X	X	X
TEMPREAS	Reasons for having a temporary job/ work contract of limited duration	X		X	X
TEMPDUR	Total duration of temporary job/ work contract of limited duration	X	X	X	X
TEMPAGCY	Contract with temporary employment agency	X		X	X

Variable	Variable label	Yearly	Quart. 1-4	Ad hoc	Special
SHIFTWK	Shift work	X		X	X
EVENWK	Evening work	X		X	X
NIGHTWK	Night work	X		X	X
SATWK	Saturday work	X		X	X
SUNWK	Sunday work	X		X	X
HWUSUAL	No. hours usually worked in main job (week)	X	X	X	X
HWACTUAL	No. hours actually worked in main job (week)	X	X	X	X
HWOVERP	Paid overtime in reference week in main job	X	X	X	X
HWOVERPU	Unpaid overtime in reference week main job	X	X	X	X
HOURREAS	Main reason for hours actually worked being different from the person's usual hours	X	X	X	X
WISHMORE	Wish to work usually more than now	X	X	X	X
WAYMORE	Way how person wants to work more hours	X		X	X
HWWISH	No. hours person would like to work in total	X	X	X	X
HOMEWK	Working at home	X		X	X
LOOKOJ	Looking for another job	X	X	X	X
LOOKREAS	Reasons for looking for another job	X		X	X
EXIST2J	Existence of more than one job or business	X	X	X	X
STAPRO2J	Professional status in the 2nd job	X	X	X	X
HWACTUA2	No. hours actually worked in 2nd job (week)	X	X	X	X
EXISTPR	Previous employment experience	X	X	X	X
YEARPR	Year in which person last worked	X	X	X	X
MONTHPR	Month in which person last worked	X	X	X	X
LEAVREAS	Main reason for leaving last job or business	X		X	X
STAPROPR	Professional status in last job	X	X	X	X
ISCOPR3D	Occupation of last job (ISCO 08, 3 digits)	X	X	X	X
SEEKWORK	Seeking employment during previous 4 weeks	X	X	X	X
SEEKREAS	Reasons for not searching an employment	X		X	X
SEEKTYPE	Type of employment sought (or found)	X	X	X	X
SEEKDUR	Duration of search for employment	X	X	X	X
METHODA	Contacted public employment office	X	X	X	X
METHODB	Contacted private employment agency	X	X	X	X
METHODC	Applied to employers directly	X	X	X	X
METHODD	Asked friends, relatives, trade unions, etc.	X	X	X	X
METHODE	Inserted/answered ads in newspapers	X	X	X	X
METHODF	Studied ads in newspapers	X	X	X	X
METHODG	Took a test, interview or examination	X	X	X	X
METHODH	Looked for land, premises or equipment	X	X	X	X
METHODI	Looked for permit, licences, financ. resources	X	X	X	X
METHODJ	Awaiting the results of application for a job	X	X	X	X
METHODK	Wait for call from public employment office	X	X	X	X
METHODL	Awaiting results of interview in public sector	X	X	X	X

Variable	Variable label	Yearly	Quart. 1-4	Ad hoc	Special
METHODM	Other method used	X	X	X	X
WANTWORK	Willingness to work not seeking employment	X	X	X	X
AVAILBLE	Availability to start working within 2 weeks	X	X	X	X
AVAIREAS	Reasons for not being available next 2 weeks	X		X	X
PRESEEK	Situation immediately before job search	X		X	X
NEEDCARE	Need for care facilities	X		X	X
REGISTER	Registration at a public employment office	X		X	X
MAINSTAT	Main labour status	X		X	X
EDUCSTAT	Student regular education in last 4 weeks	X	X	X	X
EDUCLEVL	Level of this education or training	X	X	X	X
EDUCFILD	Field of this education or training	X		X	X
COURATT	Attendance to learning activities (non-formal education) in last 4 weeks	X	X	X	X
COURLEN	No. of hours learning activities in last 4 weeks	X	X	X	X
COURPURP	Purpose of the most recent learning activity	X		X	X
COURFILD	Field of the most recent learning activity	X		X	X
COURWORH	Recent learning activity during working hour	X		X	X
HAT97LEVEL	Highest level of education or training successfully completed	X	X	X	X
HATFIELD	Highest level of education/training completed	X		X	X
HATYEAR	Year when highest level of education or training was successfully completed	X		X	X
WSTAT1Y	Situation regarding activity one year before	X		X	X
STAPRO1Y	Professional status one year before	X		X	X
ILOSTAT	ILO work status	X	X	X	X
NA11S	Economic activity in main job by sector (NACE Rev 1)	X	X	X	X
ISCO1D	Occupation (ISCO-08, 1 digit)	X	X	X	X
NA112JS	Economic activity in 2nd job by sector (NACE Rev 1)	X	X	X	X
NA11PRS	Economic activity in previous job by sector (NACE Rev 1)	X	X	X	X
ISCOPR1D	Occupation of last job (ISCO 08, 1 digit)	X	X	X	X
DURUNE	Duration of unemployment	X	X	X	X
EDUC4WN	Education or training received last 4 weeks	X	X	X	X
NA111YS	Economic activity one year before by sector (NACE Rev 1)	X		X	X
METHFW	Main method to find work last 4 weeks	X		X	X
LENTR4W	Total length of education or training	X		X	X
PURP4W	Purpose of the training received last 4 weeks	X		X	X
HWUSTRAI	Usual No. of hours of training per week	X		X	X
ISCO3D	Occupation (ISCO-08, 3 digits)	X	X	X	X
STARTIME	Time since person started to work (months)	X	X	X	X
LEAVTIME	Time since last worked (months)	X	X	X	X
LEAVCLAS	Time since last worked (classes) (months)	X	X	X	X

Variable	Variable label	Yearly	Quart. 1-4	Ad hoc	Special
INCDECIL	Monthly (take home) pay from main job (deciles)	X		X	
AGERESID	Age at which person last established their usual residence in the country	X	X	X	
HHCHILDR	Presence of children in the same household	X		X	
HHWKSTAT	Working status of adults living in household	X		X	
HHPARTNR	Presence of the partner in same household	X		X	
HHPARENT	Presence of father/mother in household	X		X	
HHCOMP	Household type (aggreg. househ. composit.)	X		X	
HHNBCHLD	No. of children in household (<25 y.), based on HHPERS code 1, 2 and 3	X		X	
HHNB0014	No. of children in household (<15 y.)	X		X	
HHNBCH2	No. of persons 0-2 years in household	X		X	
HHNBCH5	No. of persons 3-5 years in household	X		X	
HHNBCH8	No. of persons 6-8 years in household	X		X	
HHNBCH11	No. of persons 9-11 years in household	X		X	
HHNBCH14	No. of persons 12-14 years in household	X		X	
HHNBCH17	No. of children 15-17 years in household, based on HHPERS code 2	X		X	
HHNBCH24	No. of children 18-24 years in household, based on HHPERS code 2	X		X	
HHNBPERS	Total No. of persons in household	X		X	
HHNBEMPL	No. of employed adults in household (≥15 y., based on HHPERS code 4)	X		X	
HHNBINAC	No. of inactive adults in household (≥15 y., based on HHPERS code 4)	X		X	
HHNBUNEM	No. of unemployed adults in household (≥15 y., based on HHPERS code 4)	X		X	
HHNBWORK	No. of employed persons in household ≥15 y.	X		X	
HHAGEYG	Age of the youngest child in household <25 y.	X		X	
HHAGE14	Age of the youngest child in household <15 y.	X		X	
HHNBOLD	No. of persons aged 65 or older in household	X		X	
HHPERS	Definition of children and adults	X		X	
COUNTRY	Country	X	X	X	X
NATIONAL	Nationality	X	X	X	X
COUNTRYB	Country of birth	X	X	X	X
NA111D	Economic activity in main job (NACE Rev 1.1, 1 digit)	X	X	X	X
NA111Y1D	Economic activity one year before (NACE Rev 1.1, 1 digit)	X		X	X
NA11PR1D	Economic activity in previous job (NACE Rev 1.1, 1 digit)	X	X	X	X
NA112J1D	Economic activity in 2nd job (NACE Rev 1.1, 1 digit)	X	X	X	X
HATLFATH	Education level of the father (same household) (ISCED-97, 3 levels)	X		X	
HATLMOTH	Education level of the mother (same household) (ISCED-97, 3 levels)	X		X	

Variable	Variable label	Yearly	Quart. 1–4	Ad hoc	Special
HATLEV1D	Level of education (ISCED, 3 levels)	X	X	X	X
NACE1D	Economic activity in main job (NACE Rev2, 1 digit)	X	X	X	X
NACE1Y1D	Economic activity one year before (NACE Rev2, 1 digit)	X		X	X
NACEPR1D	Economic activity in previous job (NACE Rev2, 1 digit)	X	X	X	X
NACE2J1D	Economic activity in 2nd job (NACE Rev2, 1 digit)	X	X	X	X
COUNTR1Y	Country of residence one year before	X		X	X
COUNTRYW	Country of place of work	X	X	X	X
YEARESID	Years of residence in this country	X	X	X	X
REGION	Region of household (NUTS)	X	X	X	X
REGIONW	Region of place of work (NUTS)	X	X	X	X
REGION1Y	Region of residence one year before (NUTS)	X		X	X
SIGNISAL	Continuing receipt of the wage or salary	X	X	X	X