**Effects of attrition on longitudinal EU-LFS estimates**

Hannah Kiiver (hannah.kiiver@ec.europa.eu)[[1]](#footnote-1)

**Keywords:** labour market flows, longitudinal data, attrition bias

# Introduction

All countries that deliver Labour Force Survey data to Eurostat make use of a rotational pattern that results in annual overlaps of data, and with the exception of Germany, also overlaps of quarterly data. Rotational patterns are generally used in sample surveys to reduce the variances of estimators. For the EU-Labour Force Survey (LFS), this longitudinal feature can be exploited to produce highly demanded estimates of labour market transitions between the status employed, unemployed and economically inactive. When defining the longitudinal sample, Eurostat has to rely on the available overlapping data; this means, that information on individuals who drop out of the survey before completing all interviews is lost. The individuals dropping out may have moved away, it may have been impossible to contact them, they may have refused to answer, or they may have died. Eurostat does not have information on the reason why individuals drop out,nor the information that would be necessary to fill or estimate the missing labour market data for those individuals. If those individuals that drop out are on average no different from those that stay in the sample, this attrition will only lead to less precise estimates. If, however, these individuals differ with regard to their labour market status from the remaining sample, estimates based on this sample would not be representative. In short, transition statistics derived by Eurostat using LFS data may possibly suffer from attrition bias. In this paper, simple binary regression is used to determine whether attrition bias may be a problem, and simple simulations are used to estimate the potential size effect differential attrition might have on the estimates.

# Methods

There are in total 7 different rotation patterns in use in the EU-LFS at the moment, such that comparable attrition measures can only be derived for attrition after one quarter as well as after one year. Table 1 gives an overview of the current situation, excluding Germany due to the missing quarterly overlap in its rotational pattern. In this table "in" refers to a quarter in which an interview should take place, "out" to a quarter where an individual is not to be interviewed. This means that the first pattern describes a situation where individuals are to be interviewed in two consecutive quarters, than no interview takes place for two quarters, then the final two interviews are to be conducted. Following the same logic, the second pattern refers to 5 consecutive quarterly interviews.

Table 1: Overview of EU-LFS rotational patterns with quarterly overlap

|  |  |
| --- | --- |
| Rotational pattern | Country |
| 2 in 2 out 2 in | **BE,BG,DK,EE,HR,IT,LV,LT,MT,PL,RO,CH** |
| 5 in | **CZ,IE,NL,AT,SK,UK,LU** |
| 6 in  | **EL,ES,FR,CY,HU,PT** |
| 8 in | **SE,NO** |
| 3 in 1 out 2 in | **SI,FI** |
| 3 in 2 out 2 in | **IS** |

The data used covers the years 2010 to 2017. The quarterly attrition rate is defined as the share of individuals who answered the LFS in one wave and should have answered it in the next wave, but did not, expressed as % of initial answers. For this paper, quarterly samples have been matched for each year, and the results are presented for quarterly averages for each year.

Annual attrition rates refer to the share of individuals who answered the LFS in any one quarter of year t, and should have answered in the same quarter of year t+1. Again, averages of these results are taken.

In a second step, differential attrition for quarterly as well as annual data is determined by running logistic regression on the probability of staying in the sample. Since in some cases repeated observations of the same individuals are part of the sample, cluster-robust standard errors are calculated. The vector of regressors contains a number of variables that may be related to drop-out or refusals. Those regressors, or groups defined by interaction of regressors, with significantly different odds in the probability to stay in the sample are identified, and their share in the population is determined.

Two simulations are run for each of these groups: one, where the labour market status in the target quarter is estimated based on the available data, and one, where the extreme assumption of one labour market status only is made. Transition rates for both of these scenarios are calculated.

# Results

--------------------Work is still ongoing, and results are not yet finalised---------------

Results vary strongly by country, and for some countries between quarterly and annual attrition. Figure 1 shows the extent of average quarterly transition for 2016, and the average annual attrition for 2015-2016, sorted by average quarterly attrition rate.

Figure 1: Average quarterly and annual attrition rates, 2015-2016, in %

After running the logistic regressions, for some countries there is little evidence of differential attrition to begin with. Other countries show sizeable differential attrition, but the concerned population groups are so small that overall transition rates are not influenced under either scenario. For the largest group of countries though, differential attrition can be found for a sizeable group of individuals. Simulation of the two scenarios shows that there generally is an impact on the flow estimates, but that the transition probabilities calculated in each of the scenarios nevertheless rarely falls outside the confidence intervals of the original estimate.

# Conclusions

At this point in the analysis, there is evidence of some differential attrition in a number of countries. However, overall transition estimates are generally not biased in a significant way, and there is no evidence that transition estimates should or cannot be used due to attrition bias. This conclusion may differ for specific breakdowns of transition statistics, and the potential impact as well as strategies to correct for bias should be investigated.

1. European Commission, Eurostat [↑](#footnote-ref-1)