# Small area estimation of the LFS-based monthly unemployment rate in Poland

**Keywords:** *small area esitmation, space-state models, LFS, monthly unemployment rate*

## 1 Introduction

The monthly unemployment rate (MUR) is one of the main indicators published by Eurostat. It is also one of the key economic indicators used in the Macroeconomic Imbalance Procedure (MIP), the supervisory mechanism, designed to enable early detection and correction of a possible excessive macroeconomic imbalance.

Until 2008, monthly unemployment figures used to be produced by Eurostat for almost all Member States. This production was based on country-specific methods, although this approach made the production process quite burdensome. Moreover, the resulting data were only comparable to a very limited degree. In 2008, Eurostat developed a common data estimation method based on quarterly data from the labour force survey (LFS) in accordance with the definition of the International Labour Organization (ILO) and monthly registered unemployment figures. The method was adopted by all countries, except those which had already produced monthly data themselves. However, according to Eurostat, the system is still far from ideal. From their point of view, the Member States are in a better position to compile national statistics as they have access to all available data sources. They are also able to better assess the quality of these sources, hence the quality of results, and on this basis decide which method of data production is best suited for their country. Therefore, and also because of the ongoing work regarding the Integrated European Social Statistics Framework Regulation (IESS FR), which will also regulate the MUR issue, Eurostat encourages the Member States to test their MUR production methods and to gradually take over the production of monthly rates of unemployment based on the ILO definition.

The general objective of the paper is to present the project aimed to develop the method for estimating the monthly unemployment rate in Poland based on the ILO methodology using small area estimation methods. The following specific goals were set to facilitate the achievement of the main goal:

- review methods of measuring unemployment,
- compile LFS microdata for the indirect estimation of the monthly unemployment rate,
- assess data on registered unemployment in terms of their usefulness for MUR estimation,
- select a method for the indirect estimation of MUR.

## 2 Methods

In the project we tested several approaches including small area estimation methods that account for quarterly time series and characteristics of the Labour Force Survey (e.g. rotation panel-design survey).

The idea of using structural time series (STS) models for indirect estimation involves constructing a model for a series of direct estimates $\hat{Y}_{d,t}$ $(d = 1, \ldots, D,\ t = 1, \ldots, n)$ in order to remove variability due to random sampling. In addition to accounting for unobservable components typically included in structural time series models, the model in question is also designed to account for the component representing the error due to direct estimation. The direct estimator $\hat{Y}_{d,t}$ can be expressed as a sum of the true value of the estimated parameter $Y_{d,t}$ and the estimation error $e_{d,t}$:

$$\hat{Y}_{d,t} = Y_{d,t} + e_{d,t}. \tag{1}$$

In order to determine the true value of the estimated parameter and the estimation error from a series of direct estimators, the first of these components is expressed as an STS model. If a basic structural model is used, the direct estimator is given by:

$$\hat{Y}_{d,t} = L_{d,t} + S_{d,t} + I_{d,t} + e_{d,t}. \tag{2}$$

The estimation error in periodical surveys has a variance which varies over time, as a result of changes in sample size, methodological changes or changing levels of the phenomenon of interest. When a survey is based on a rotating panel design, estimation errors are also characterised by autocorrelation. These two characteristics of estimation errors – heteroscedasticity and autocorrelation – can be accounted for by expressing this component as:

$$e_{d,t} = \sigma_{d,t} e'_{d,t}, \tag{3}$$

where $\sigma^2_{\hat{Y},d,t}$ denotes the variance of estimator $\hat{Y}_{d,t}$, whereas $e'_{d,t}$ denotes an ARMA($p$,$q$) process with zero mean and unit variance.

The choice of the model formula, which depends on the selection of appropriate values of $p$ and $q$ is closely connected with the use of the rotating panel design, and its parameters, in practice, are chosen on the basis of estimated survey error autocorrelation coefficients. Pfeffermann et al. [1] proposed a method of estimating autocorrelation coefficients in the case of a rotation panel-design survey.

By adopting an indirect estimation approach, which involves the use of STS models, it is possible to borrow strength from other domains. If, for instance, changes in the slope of the trend line in domains $d_1$ and $d_2$ display similar patterns, one can use a multivariate model in which variance between residuals $\eta_{R,d,t}$ components $R_{d,t}$ $(d = d_1, d_2)$ is non-zero:

$$Cov(\eta_{R,d,t}, \eta_{R,d',t'}) = \begin{cases} \sigma^2_{R,d,t} & d = d', t = t' \\ \zeta_{R,d,d',t} & d \neq d', t = t' \\ 0 & t \neq t', \end{cases} \tag{4}$$

where $\zeta_{R,d_1,d_2,t} > 0$.

This practice can make the model more resistant to outliers, which can cause considerable deviations from the correct trend line, above all, if they are present at the end of the series.

To increase the accuracy of the estimation, the basic structural model can be extended to include a component that accounts for the influence of explanatory variables $\mathbf{X}_{d,t} = (X_{d,1,t}, \ldots X_{d,K,t})'$ on the value of the dependent variable $Y_{d,t}$, by adding a regression

component ${}_{d,t}\mathbf{X}_{d,t}$. Another possibility to use auxiliary variables is to model $\mathbf{y}$ and $\mathbf{X}$ by a multivariate structural time series model and include the correlation between slopes or residuals. In the second case, it is assumed that external factors – not described by trend and seasonality – influence in a similar way both $\mathbf{y}$ and $\mathbf{X}$. For more details, see for example [2, 3].

## 3 Results

For the purpose of estimation monthly unemployment rate in Poland different models were considered, for instance depending on auxiliary variables.

Gains in precision were especially visible for men and women aged 25+, where CVs are small (maximum 5.3% in the case of model 5) and results are acceptable. For men and women aged 15–24, CVs range from 4 to 9.99% in the case of model 3. In general, the STS models were found to be useful for estimating MUR.

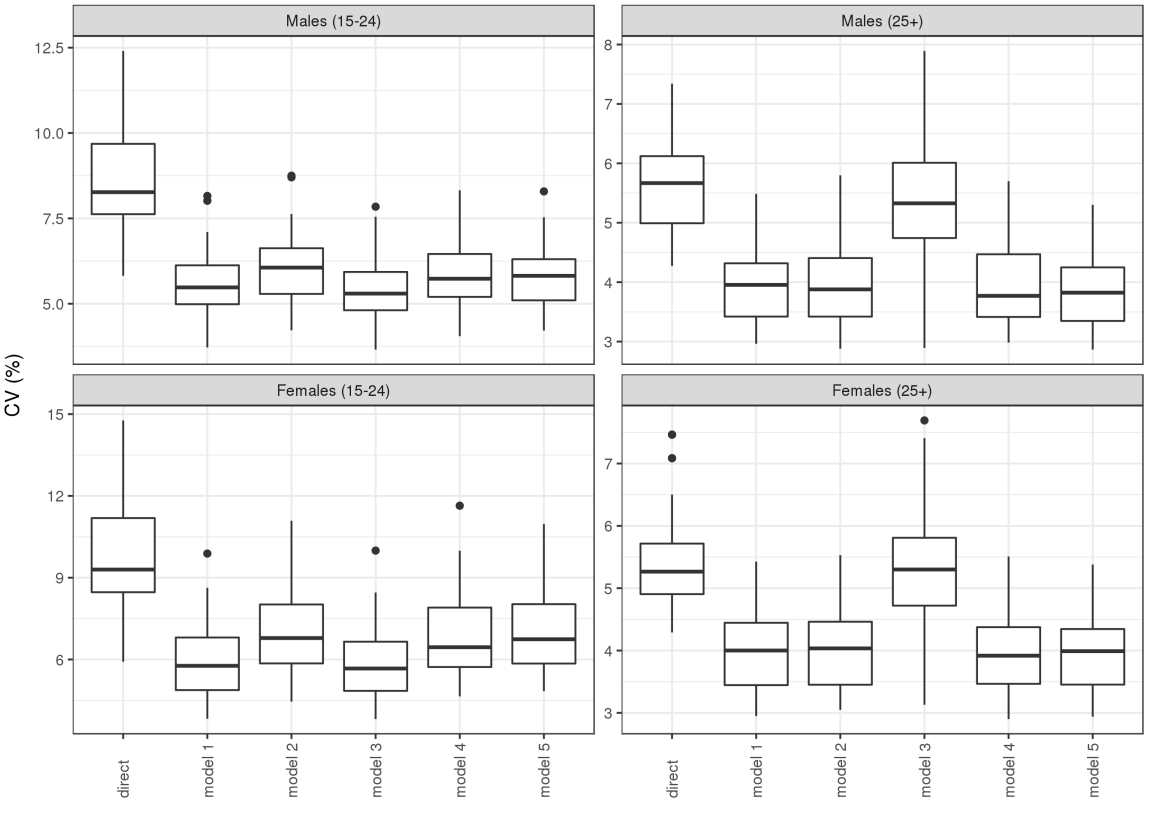


Figure 1: Coefficients of variation of direct and indirect estimators of monthly unemployment rate

## 4 Conclusions

The main goal of the project was to develop a method for estimating the monthly unemployment rate based on the ILO methodology using indirect estimation techniques.

Conclusions presented below are the selected results from the completed research work, which will, hopefully, be used to inform decisions regarding the future production of monthly unemployment rate statistics for Poland.

Some chosen conclusions:

- as confirmed by results obtained in the project, it is not possible to calculate the ILO monthly unemployment rate based only on the LFS owing to insufficient LFS sample size;
- the precision of indirect estimates of the monthly unemployment rate could be improved by calibrating design weights. The results indicate that the two-step calibration approach can serve as the basis for indirect estimation;
- it was also found that the use of STS models significantly improves precision compared with direct estimation.

## Acknowledgements

The paper is based on "Quality improvement of the monthly unemployment rate – MUR (grant agreement No. 07131.2016.002-2016.326 of 11.08.2016).

## References

[1] D Pfeffermann, M Feder, and D Signorelli. Estimation of Autocorrelations of Survey Errors With Application to Trend Estimation in Small Areas. *Journal of Business & Economic Statistics*, 16(3):339–348, 1998.

[2] Jan A Van den Brakel and Sabine Krieg. Estimation of the monthly unemployment rate through structural time series modelling in a rotating panel design. *Survey Methodology*, 35(2):177–190, 2009.

[3] Jan A van den Brakel and Sabine Krieg. Dealing with small sample sizes, rotation group bias and discontinuities in a rotating panel design. *Survey Methodology*, 41 (2):267–296, 2015.