

LITHUANIAN LABOUR FORCE SURVEY: ANALYSIS OF BREAKS IN TIME SERIES DUE TO THE INTRODUCTION OF THE IESS

Introduction

The adoption of the Integrated European Social Statistics (IESS) Framework Regulation (EU) 2019/1700 and Implementing Regulation (EU) 2019/2240 brought significant changes to the whole process of conducting the Labour Force Survey. The amendments of definitions and classification of employed and unemployed might cause breaks in time series. Therefore, this report aims to briefly present the analysis of the break in time series due to the introduction of the IESS FR and draw the conclusions from the results.

Overview of the main changes in the LFS

First of all, the definition of employed was modified, concerning the classification of people with a job who were absent during the reference week. Following the new IESS FR, people absent for other reasons, such as lay-offs, are counted as employed only if the expected period of absence from work does not exceed 3 months. Otherwise, people who are temporarily unemployed on a full-time basis for more than three months, are no longer classified as employed. Also, people on parental leave are counted as employed only if they receive or have a right to receive work-related income or (parental) benefits or were to remain on child care leave presumably for less than three months. Furthermore, seasonal workers are classified as employed if during the off season the worker continues to regularly perform work related tasks and responsibilities.

Another change was made regarding persons producing agricultural goods. If persons produce them exclusively or mainly for self-consumption and simultaneously do not have any other job, they are excluded from employment category.

As of unemployment classification, only active methods of searching a job are included. For the purposes of identifying active job search, such activities are: studying job advertisements; placing or answering job advertisements; placing or updating CVs online; contacting employers directly; asking friends, relatives or acquaintances; contacting a public employment service; contacting a private employment agency; taking a test, interview or examination as part of a recruitment process and making preparations to set up a business.

Statistical analysis of changes in Lithuanian LFS

Due to the changes in definitions for employment and unemployment, it is crucial to evaluate the impact on Lithuanian LFS data.

The analysed data (time series of employed and unemployed) consists of quarterly estimates from 1998 Q1 to 2021 Q3. In this case time series consists of 87 observations and each of observation corresponds to particular data time point. If there is any kind of break down in time series it is expected to observe time series change point in 2020 Q4 (change point 84) or in 2021 Q1 (change point 85).

In order to evaluate possible break in time series, various R packages were used:

- Mcp: Regression with Multiple Change Points
- EnvCpt: Detection of Structural Changes in Climate and Environment Time Series
- Segmented: Regression Models with Break-Points / Change-Points Estimation
- Strucchange: Testing, Monitoring, and Dating Structural Changes

- Cpm: Sequential and Batch Change Detection Using Parametric and Nonparametric Methods
- ChangePoint: Methods for Changepoint Detection
- Bcp: Bayesian Analysis of Change Point Problems
- TSMCP: Fast Two Stage Multiple Change Point Detection

First, we manually mark possible change points in time series. It is possible that there is a change point at time point from 4 to 10 and from 34 to 37. We additionally mark time point 84 as we expect to detect change in time series and confirm time series break at the time period 2020 Q4 or 2021 Q1. By manually setting the structure of time series we are looking for three change point between four segments.

R package mcp

At first, we plot estimates of total employed. We construct simple intercept only model for detection of change point using R package mcp. The summary shows good parameter recovery, it obvious that the change point is detected around time point 36, but there is no indication to have time series break at 84. Plotting the posterior distributions of the change points reveal that they are not well represented by a Gaussian or other known distributions.

Next, we apply simple autoregressive model. The most common use case is probably just to add AR(1) to the first segment. The results are quite similar to intercept model, clear change point detected at time point near 36, but as in previous example no identification of change point at 84.

R package EnvCpt

Another method used is R package EnvCpt which can detect change points in mean and variance (not separately), slopes, and AR(1)/AR(2), as well as conveniently fitting various models without change points. It automatically infers the number of change points. Unless otherwise instructed (through models argument), EnvCpt fits all models to the data, allowing you to pick one. We get maximum-likelihood estimates of the change points at 36; 61; 87 and the parameters of each segment. However, the change point at 87 should just be ignored as this is the last observation in time series. Thus, this model once again do not indicate the break in time series at point of our interest.

R package Segmented

R package Segmented is one of the most popular package for change point analysis. It has a very shallow learning curve combined with great modelling flexibility. The disadvantage of package segmented is that you can only have one kind of segments and (for some reason) it ignores the intercepts on anything but the first segment. We build a simple linear intercept-only model. There is no change point at 85.

R package strucchange

Another option is to apply intercept-only model using strucchange package function breakpoints. The difference is that it is limited to gaussian residuals, and it actually models intercepts, quadratic terms, etc. It scans through fits with 1, 2, 3, . . . N break points and determine where the optimal break points between this number of segments would be. It corectly identifies change point at time period 36.

R package cpm

R package cpm is an intercept-only (in mean and variance) package, so it cannot model slopes. It can detect single change points via detectChangePoint and multiple change points via processStream. processStream is an automatic change point detection, using a p-value threshold to determine if a candidate should be marked as a hit. The detected change points are shown below.

```
## [1] 36 57 64
```

Thus, the change points detected do not indicate the break at change point of our interest.

R package changepoint

R package changepoint is focused on intercept-only changes. It can estimate changes in means, variance, or both. It is semi-automatic in that you can set the number of change points using parameter Q and this defaults to five. It can recover ML estimates of the intercepts. It does not estimate uncertainty, nor model checking. It only takes a response variable, so the change point is the data index, not the point on an x-axis. The data should be ordered. In our case it is ordered and we have 1 data point at each x. The results show that only one change point at time point 36 is detected. The package changepoint.np extends changepoint by providing a non-parametric version, which returns change points:

```
## [1] 3 9 36 57 87
```

R package bcp

R package bcp automatically detects change points and segment types. It provides estimates of means and probability of change point at each x-coordinate. It has little additional functionality. The results are similar to application of changepoint package and the change point captured is 36.

R package TSMCP

The R package TSMCP refers to Time-Series Multiple Change Point. It detects two change points automatically:

```
## [1] 36 58
```

Therefore, the change points detected do not indicate the break at 2020 Q4 or 2021 Q1.

To sum up all the methods used, the analysis shows that change point in time series is detected at time point 36 (2008Q4). It might be the situation that applied methods can not detect more change points, especially we are interested in (2020Q4 or 2021Q1). For this reason, we are reducing the length of our time series and we are taking into analysis time points from 2012Q1 (time point 49) to 2021Q3 (time point 87).

The intercept-only model does not detect any kind of changes in time series at the interested time points. Additionally, we apply autoregressive model to time series. The results are the same, there is no evidence that we are observing time series break at the interested time point.

The same methods were used to check if there is the break in time series of estimates of total unemployment. Furthermore, the estimates of total employed by age groups 15-24, 25-54, 55-64 and 65+ were calculated, as well as total unemployed of age groups 15-24, 25-54 and 55-64.

In conclusion, none of the applied methods detects break in time series of estimates at time point 84 or 85.

Further analysis of the results

As concluded from the statistical analysis, the applied methods to detect change points (breaks) in time series of estimates of total employed and unemployed do not indicate that there is any kind of break in time series for particular time point (2021 Q1) due to the changes of IESS FR. However, the further question arises – what could be the possible reasons why there were no breaks detected at time point of 2021 Q1.

Parental leave

The very beginning of this document generally overviewed the main changes in the LFS. In terms of employment definition, some changes were made regarding to persons on parental leave. The definition of employed included persons who were not at work during the reference week but had a job or business from

which they were temporarily absent. In terms of parental leave there was no criterion of the duration of absence or the job-related income or benefits. In case of Lithuania, the Labour Code establishes parental leave as one of the target types of leave during which the employment relationship between the employer and the employee continues throughout the leave period. That means persons on parental leave were classified as employed because of continuous employment relationship.

The new IESS introduced several additional conditions defining employment. People on parental leave are counted as employed only if they receive or have a right to receive work-related income or (parental) benefits or were to remain on child care leave presumably for less than three months. According to the national legislations of Lithuania, parental leave can last up to the child is three years old, but parental benefit is paid for either the child is up to one year old or two years old, at the choice of the recipient. That could suggest a possible increase in economically inactive persons group because of people on parental leave on their third year when no benefits are paid.

However, the majority of people who indicated the reason of absence of work as parental leave specified receiving the benefit of 50 per cent or more of their wage. Also, as there are no parental benefits paid on the third year at all, it is common to enroll children to early childhood education and come back to work. Thus, for these reasons it is possible that the change of IESS did not have a significant impact on labour force indicators.

Lay-offs

Another change was made regarding temporary absence, for instance, lay-offs. Pre-IESS people on lay-offs were classified as employed if they had an assurance of return to work within a period of 3 months or received 50 per cent of their wage from their employer. Now the decision to categorise a person as employed is based only on the expected duration of the absence.

Cases of lay-off before Covid-19 pandemic were very low in Lithuania: according to the LFS data on average less than 1 per cent indicated the reason of absence being on the lay-off in 2019. Due to the pandemic, the strict quarantine started in the mid March of 2020 and the percentage of lay-off as the reason of absence increased to 4.8 per cent in 2020 Q1 and to 11.5 per cent in 2020 Q2. The majority of the respondents who were on lay-offs in 2020 Q2 answered that they are planning to come back to work in 3 months or less and only a few respondents answered no and received 50 per cent or more of their wages, thus they were counted as employed. That implies that in most cases the employees are planning to come back to work in 3 months or less and do not necessarily receive 50 per cent or more of the wage. Also State Labour Inspectorate establishes that employee must be informed how long the lay-off is going to last (the start and the end date). The employer has to pay at least the minimum monthly salary if the employee works full time or less accordingly. Therefore, the law sets the minimum boundary of the payment in accordance with the minimum wage rather to usual employee's wage. For this reason, it is likely that the elimination of 50 per cent or more wage criterion does not affect the data significantly. The criterion of the duration seemed more substantial even before the new IESS, as the majority of employees referred to planning coming back to work in 3 months or less.

Seasonal work

According to the new IESS, they are classified as employed if they continue to regularly perform tasks and duties during the off-season for the job or business. Previously seasonal workers were classified as employed if they had the assurance to come back to work with the same employer at the beginning of the next season and the employer continued to pay at least 50 per cent of their wage or salary during the off-season. It is impossible to evaluate the effect of the change, as the criteria have changed: performing work vs. continuous receipt of a salary. Even though the classification has changed since the new IESS came into force, it is unlikely

it could have had relevant impact on labour force indicators because seasonal work in Lithuania does not contribute a lot to the labour market situation.

Conclusion

The analysis of estimates of total employed and unemployed and estimates by age groups do not indicate that there is any kind of break in time series for interested time periods because of the changes in IESS. As a result, the Lithuanian LFS does not implement any break-correction.