

Tax Compliance in Romania

Monica Robayo-Abril

Georgiana Balaban

Marcin Wronski

REPLICATION PACKAGE

This package replicates estimates reported in the paper. The package is divided into two parts:

- 1) Imputation of tax income into EU-SILC data set (Romania, 2021 EU-SILC)
- 2) Calculation of estimates presented in the paper.

The first step is time-intensive and requires enough computing power (the imputation is based on 6 million tax records). Therefore, we also provide a ready, imputed EU-SILC dataset, which allows you to replicate estimates even if your PC is not ready to handle the imputation.

Data availability statement

Tax administrative data

- Tax administrative data is strictly confidential and provided by the Romanian Ministry of Finance, as a result of a strong collaboration. There is no standard procedure to request the data, as the data is not made publicly available

EU-SILC microdata

- Access to EU-SILC, EU-LFS, and the harmonized HBS from Eurostat is based on having gone through an approval process by Eurostat that includes information on the research to be undertaken and the products emerging, a list of users with access to the data, a dedicated data and research manager.
- Access will be granted to all those with a valid research agreement, and will be removed at the end of the agreement unless the user sends an extension or revised agreement.
- More details on microdata access can be found here
- <https://ec.europa.eu/eurostat/web/microdata>

The structure of the replication package

1. Data:

- a) Raw tax data (Tax_07_2020) – this file is used for a few calculations
- b) Tax data prepared for the imputation (Tax_07_2020_for_imputation) – this file is the basis for imputation
- c) ROU_SILC_2021 – this file is based on the imputation

If you run the Imputation code, the fourth file (imputed EU-SILC) will be added.

Due to legal constraints, this folder is not shared as a part of the replication package.

2. Data ready

Because some PCs may not be able to handle the imputation, we also include the imputed file in the replication package.

3. Code

- a) Imputation of tax income into EU-SILC.do

This file imputes tax income into the EU-SILC data set. It's time-intensive.

- b) Tables and Figures

This file runs the calculations needed to produce tables and figures.

As you do not have access to original tax files you need to omit the relevant parts.

4. Log

This folder includes logs from the execution of the do files.

5. Master Excel with Tables

Here you can find the results presented in the paper. You can also check the manual calculation of compliance by sectors (Table 6) and by age-gender cells (Table 7, Table 8). The compliance rates were calculated based on the median values of survey income and imputed tax income across sectors and age-gender cells. To replicate the tables you need to manually input the relevant estimates into Excel.

6. Paper

This file includes the paper in pdf and docx format.