

README for the Reproducibility Package for Statistical Matching for Combining the European Survey on Income and Living Conditions and the Household Budget Surveys: An Evaluation of Energy Expenditures in Bulgaria

This reproducibility package presents code and documentation to reproduce the paper "Statistical Matching for Combining the European Survey on Income and Living Conditions and the Household Budget Surveys: An Evaluation of Energy Expenditures in Bulgaria" by Britta Rude and Monica Robayo-Abril. The authors prepared the research piece as part of the EU Poverty Program of the Poverty and Equity Global Practice of the World Bank in 2023. We thank Eurostat for providing us with the necessary data for this project. We also thank Sergio Oliviera for valuable comments and suggestions.

Outline and Folder Structure

The repository consists of the following folders:

1. **Data** has related information about the underlying data. The authors do not have permission to share the data. However, interested researchers can request them from Eurostat and the Bulgarian National Statistics Institute and use them under strict data usage agreements.
2. **Code/Do** has the related codes. Researchers who wish to reproduce the work conducted must adjust the global path in line 26 of the master do-file to reflect their own file structure.
3. **Code/Results** stores all the related output after the code runs: figures and tables.

How to Reproduce

1. **Software Requirements:** Ensure you have Stata installed on your machine.
2. **Request the data:** The package relies on two different data sources which have to be requested from government agencies:
 - a. The EU-SILC can be requested from Eurostat under a data usage agreement. For more information on how to get access to the EU-LFS see this link. For more information on how to request access to the EU-SILC see this link: <https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions>. EU-SILC data files should be stored in the folder "Data/".
 - b. The HBS 2019 can be requested from the National Statistical Office in Bulgaria. To initiate the data access process, the researcher needs to formally request access to the anonymized microdata conducted by the Bulgaria National Statistical Institute (NSI). The request, addressed to Desislava Dimitrova, Head of Statistics in the Living Conditions Department, needs to be submitted in the form of a letter outlining the specific research projects, a detailed research proposal, confidentiality agreements signed by each member of the team, and some standard forms sent by the NSI. HBS 2019 data files should be stored in the folder "Data/Raw/".
3. **Download the reproducibility package:** Download the compressed reproducibility package to your local machine.

4. **Run the Code:** Execute the Stata code by running the master do-file located in the “Code/Do” directory. Before that, adapt the path in line23 of the master do-file.
5. **Explore Results:** Review the generated output and figures in the folder “Code/Results”.

Feedback and Suggestions

Research is an ongoing process, and while we have made every effort to ensure accuracy, errors may exist. We appreciate your understanding and are committed to addressing any identified issues to enhance the overall quality of the research.

Thank you for exploring this project. For questions, comments, and suggestions please contact: Britta Rude, Young Professional, brude@worldbank.org and/or Monica Robayo-Abril, Senior Economist, mrobayo@worldbank.org from the Poverty and Equity Global Practice at the World Bank.