



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

Second Submission: RR_BGR_2024_93

Luis Eduardo San Martin

PRWP Reproducibility Verification

reproducibility@worldbank.org

May 24, 2024

This review verifies the reproducibility of the exhibits included in 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*”.

Contents in this review:

1. Main findings
2. List of exhibits and reproducibility status
3. Reproduction Environment

Main findings

- The code was successfully executed on a new computer after:
 1. Changing the file path globals.
 2. Specifying the bin width in line 110 of do-file *01_01_Matching_Harmonize_HBS.do* so that the width is not set by default in Stata.
- The output demonstrates consistent stability across multiple runs. Specifically, executing the code two times consecutively yielded identical results.
- The code takes approximately 20 minutes to run.
- We conducted our reproducibility analysis based on the paper uploaded to the GitHub repository with the code on May 17, 2024.
- Every exhibit has been reproduced accurately.
- **Reproducibility Summary:**
 - **Data:** Data is confidential and has not been included in the reproducibility package. For more details, please refer to the README file.
 - **Code:** All code files (from cleaning to analysis) are included in the reproducibility package.
 - **Outputs:** All outputs are generated by code included in the reproducibility package.

- **Reproducibility verification:** Reviewers used data provided directly by the authors to conduct the reproducibility verification, and this is not included in the public reproducibility package.
- **Dependencies environment:** The reviewers created a new environment for dependencies using the latest versions available for each dependency at the moment of the review.

List of exhibits and reproducibility status

Results in the Main Section of the Paper

- **Table 1** Does not apply: the exhibit does not show analysis results
- **Table 2** Reproduced. The exhibit was compared against *Summary_energyvars.doc*
- **Figure 1** Reproduced. The exhibit was compared against *Energyspendinshare_quintiles.xlsx*
- **Table 3** Reproduced. The exhibit was compared against *summary_missing.doc*
- **Figure 2** Reproduced. The exhibits was compared against *kdensity_hbs_eusilc_regress1_restricted.png*
- **Figure 3** Reproduced. The exhibits was compared against *kdensity_hbs_eusilc_pmm1_weighted.png*
- **Figure 4** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_truncated3.png*
- **Figure 5** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm2_female_weighted.png*
- **Figure 6** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm2_male.png*
- **Figure 7** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm2_highincome.png*
- **Figure 8** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm2_lowincome_weighted.png*
- **Figure 9** Reproduced. The exhibit was compared against *eshare_comparison_quintiles.png*
- **Figure 10** Reproduced. The exhibit was compared against *poverty_overlap.png*

Results in the Appendix

- **Figure A 1** Reproduced. The exhibit was compared against *hist_income_equivalized_eusilc.png*
- **Figure A 2** Reproduced. The exhibit was compared against *hist_income_equivalized_hbs.png*
- **Table A 1** Does not apply: the exhibit does not show analysis results
- **Figure A 3** Reproduced. The exhibit was compared against *hist_energyspendingshare_hbs_2019.png*
- **Table A 2** Does not apply: the exhibit does not show analysis results
- **Table A 3** Reproduced. The exhibit was compared against *correlationtable.rtf*
- **Figure A 4** Reproduced. The exhibit was compared against *poverty_overlap_venn.png*
- **Figure A 5** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_regress1.png*
- **Figure A 6** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_regress1_restricted.png*

- **Figure A 7** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_regress2.png*
- **Figure A 8** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_regress2_restricted.png*
- **Figure A 9** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_regress3.png*
- **Figure A 10** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_regress3_restricted.png*
- **Figure A 11** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm1.png*
- **Figure A 12** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm1_lowincome.png*
- **Figure A 13** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm2.png*
- **Figure A 14** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm2_lowincome.png*
- **Figure A 15** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm3.png*
- **Figure A 16** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pmm3_lowincome.png*
- **Figure A 17** Reproduced. The exhibit was compared against *knn_robustness_kdensity_hbs_eusilc_pmm2.png*
- **Figure A 18** Reproduced. The exhibit was compared against *knn_robustness_kdensity_hbs_eusilc_pmm2_lowincome.png*
- **Figure A 19** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_truncated1.png*
- **Figure A 20** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_truncated1_lowincome.png*
- **Figure A 21** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_truncated2.png*
- **Figure A 22** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_truncated2_lowincome.png*
- **Figure A 23** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_truncated3.png*
- **Figure A 24** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_truncated3_lowincome.png*
- **Table A 4** Reproduced. The exhibit was compared against *Imputation_energyspending_pmm2.xls*
- **Figure A 25 - Distribution Household with elderly (>64)** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_elderly_weighted.png*
- **Figure A 25 - Distribution Household with pensioner** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_pensioner_weighted.png*
- **Figure A 25 - Distribution Female-headed household** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_femalehead_weighted.png*
- **Figure A 25 - Distribution No with secondary educ.** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_secondaryeduc_weighted.png*
- **Figure A 25 - Distribution Income cat. (1-20)** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_incomecat_weighted.png*
- **Figure A 25 - Distribution Household with unemployed** Reproduced. The exhibit was compared against *kdensity_hbs_eusilc_unemployed_weighted.png*

- **Figure A 25 - Distribution Urban Reproduced.** The exhibit was compared against *kdensity_hbs_eusilc_urban_weighted.png*
- **Figure A 26 - Mean energy expenditure share by Female-headed household Reproduced.** The exhibit was compared against *eshare_femalehead_weighted.png*
- **Figure A 26 - Mean energy expenditure share by Household with pensioner Reproduced.** The exhibit was compared against *eshare_pensioner_weighted.png*
- **Figure A 26 - Mean energy expenditure share by Household with elderly (>64) Reproduced.** The exhibit was compared against *eshare_elderly_weighted.png*
- **Figure A 26 - Mean energy expenditure share by No with secondary educ. Reproduced.** The exhibit was compared against *eshare_secondaryeduc_weighted.png*
- **Figure A 26 - Mean energy expenditure share by Income cat. (1-20) Reproduced.** The exhibit was compared against *eshare_incomecat_weighted.png*
- **Figure A 26 - Mean energy expenditure share by Household with unemployed Reproduced.** The exhibit was compared against *eshare_unemployed_weighted.png*
- **Figure A 26 - Mean energy expenditure share by Urban Reproduced.** The exhibit was compared against *eshare_urban_weighted.png*
- **Figure A 27 Reproduced.** The exhibit was compared against *Energyspendinshare_quintiles.xlsx*, tab *OLS_Spec3*
- **Figure A 28 Reproduced.** The exhibit was compared against *Energyspendinshare_quintiles.xlsx*, tab *Truncreg_Spec3*
- **Figure A 29 Reproduced.** The exhibit was compared against *eshare_members_weighted.png*
- **Table A 5 Reproduced.** The exhibit was compared against *Reg1.doc*
- **Table A 6 Reproduced.** The exhibit was compared against *TruncReg1.doc*

Reproduction Environment

Paper exhibits were reproduced in a computer with the following specifications:

- OS: Windows 11 Enterprise
- 11th Gen Intel(R) Core(TM) i5-1145G7 @ 2.60GHz 1.50 GHz
- Memory available: 15.7 GB
- Software version: Stata 18 MP