

Authors: Christopher Hoy, Yeon Soo Kim, Minh Cong Nguyen, Mariano Sosa, Sailesh Tiwari

Academic Data Use

Reproducibility package for Policy Research Working Paper 10615 “Attitudes Towards Reducing Fossil Fuel Subsidies: Evidence across 12 Middle-Income Countries”

Overview

The code and data in this package will constructs the analysis files, create tables and figures for Hoy, Christopher; Kim, Yeon Soo; Nguyen, Minh; Sosa, Mariano; Tiwari, Sailesh. (2023) using Stata and R. The main Stata dofile runs all of the code in sequences to produce 19 Tables and 33 Figures in the paper. Then, the second R code will run to produce 4 word-cloud figures. The replicator should expect the codes to run for about 50 minutes.

The reproducibility package reproduces all analytical exhibits except for Tables 1 and 2. Table 1 was created manually and Table 2 uses WB-internal package dependencies that are not included in the reproducibility package. The do-files that produces Table 2 is included (“0 - GMD input.do”) but it is not called by the main do-file. The package also does not use raw data inputs but constructed data, raw data inputs are restricted and cannot be made public with the reproducibility package. Do-files for data cleaning and construction that produce the constructed data are included (“1a - Clean raw data.do” and “1b - Add weights.do”) but they are not called by the main do-file.

Summary of Availability

- All data **are** publicly available.
- Some data **cannot be made** publicly available.
- No data can be made** publicly available.

Details on each Data source

| Data.Name | Data.Files | Location | Provide | Citation |
|-------------|------------------------------|------------|---------|--|
| GMD surveys | GMD_input.dta | Data\Raw | TRUE | World Bank (2023) |
| GMD/WDI | Population char data.dta | Data\Raw | TRUE | Authors |
| Survey data | allrawdata_weight.dta | Data\Clean | TRUE | “Hoy, Christopher; Kim, Yeon Soo; Nguyen, Minh; Sosa, Mariano; Tiwari, Sailesh. 2023. Building Public Support for Reducing Fossil Fuel Subsidies: Evidence across 12 Middle-Income Countries. Policy Research Working Papers; 10615. © World Bank, Washington, DC. http://hdl.handle.net/10986/40659 |
| Survey data | q31 theme classification.dta | Data\Raw | TRUE | |
| Survey data | Bolivia_eng.xlsx | Data\Raw | TRUE | |
| Survey data | Ecuador_eng.xlsx | Data\Raw | TRUE | |
| Survey data | Egypt_eng.xlsx | Data\Raw | TRUE | |
| Survey data | Indonesia_eng.xlsx | Data\Raw | TRUE | |
| Survey data | | Data\Raw | TRUE | |

| | | | |
|-------------|-------------------------------|----------|------|
| Survey data | kazakh gsg3 surveys (1).xlsx | Data\Raw | TRUE |
| Survey data | Russian gsg3 surveys (1).xlsx | Data\Raw | TRUE |
| Survey data | Vietnam_eng.xlsx | Data\Raw | TRUE |
| Survey data | Angola_eng.xlsx | Data\Raw | TRUE |
| Survey data | Argentina_eng.xlsx | Data\Raw | TRUE |
| Survey data | Open ended Qs.xls | Data\Raw | TRUE |
| Survey data | Pakistan_eng.xlsx | Data\Raw | TRUE |
| Survey data | Bangladesh_eng.xlsx | Data\Raw | TRUE |

The authors implemented and collected survey data in 12 middle-income countries. See the paper on the survey data description and a sample questionnaire. This is the main source of data for most tables and figures.

Part of one table and two figures (population characteristics) used output from household surveys in the Global Monitoring Database using Datalibweb. The list of 12 surveys is listed in the file GMD_input.dta.

Computational requirements

Software Requirements

- Stata 17, with external programs: reghdfe, esttab, grc1leg, coefplot, cibarc, ietoolkit, ereplace, grc1leg. All of these external programs are included in the reproducibility package.
- R 4.2.1, with external packages: rmarkdown, knitr, ggwordcloud, dplyr, RColorBrewer, wordcloud, wordcloud2, htmlwidgets, textplot, quanteda.textplots, webshot2, chromote. These external packages are not included in the reproducibility package but metadata files to install them in a working environment using the package “renv” are included in the reproducibility package. Instructions to reproduce the environment are provided below.

Memory and Runtime Requirements

A standard desktop or laptop of 2022 can run the programs.

The code was last run on a 4-core Intel-based laptop with Windows version 10.14.4. All of the code were last run on a Intel(R) Core(TM) i5-10310U CPU @ 1.70GHz 2.21 GHz, with 16GB of RAM. Computation took 50 minutes.

Description of programs/code

- The dofile “master dofile.do” will run all dofile to produce tables and figures in the paper. Output files are called appropriate names (table5.tex, figure12.png) and should be easy to correlate with the manuscript.
- These three dofiles “0 - GMD input.do”, “1a - Clean raw data.do”, “1b - Add weights.do” are provided for reference. These dofile prepare the clean data for the analysis.
- Tables and figures in the main body of the paper come from the dofiles “2 - 2312_Body_v1.do” and “3 - basic stat WP paper.do”. Appendix tables and figures are from “2 - 2312_Appendix (tables)_v1.do” and “2 - 2312_Table3_FigureC8_TableC1_v1.do”. Pre analysis figures come from the dofile “5 - Pre-analysis_c1_c2_c3.do”.
- Open text tables and figures in Appendix D are from these sets “4 - 1. Cleaning and Appending.do” (to prepare the text data), “4 - 2. Identify keywords.do” (to identify theme and keywords), “4 - 3. Analysis.do” and “Wordclouds.R” (to perform analysis, keyness, and wordcloud).

Instructions to Replicators

- Define the main path of the work, and copy the folders from the reproducibility package as follows:

```
Data\  
---Clean\  
---Intermediate\  
---Raw\  
---Pre_Analysis\  
Dofile\  
Figures\  
---CSV\  
---HTML\  
Tables\  

```

- Change the path in the “master dofile.do” – line 34.
- The main dofile “master dofile.do” will run all dofiles in sequence to produce tables and figures in the paper. Output files are called appropriate names (table5.tex, figure12.png) and should be easy to correlate with the manuscript.
- Once the main dofile is done in Stata, open RR_MIC_2024_85.Rproj and follow the instructions to reproduce the R environment and run the code
- **Reproducing the R environment:** with the RStudio project RR_MIC_2024_85.Rproj active, open the script Dofile/Wordclouds.R and run lines 3-6 to install all package dependencies and reproduce the

environment of this project. They should only be run the first time the code runs, replicators should comment the lines out once the package dependencies are installed.

- **Running the R code:** with the RStudio project RR_MIC_2024_85.Rproj active, run the script Dofile/Wordclouds.R to create the word-cloud figures.

List of tables and programs

The provided code reproduces:

- All numbers provided in text in the paper
- All tables and figures in the paper
- Selected tables and figures in the paper, as explained and justified below.

Main body of the paper

| Figure/Table # | Program | Line number | Output file | Note |
|----------------|---|-------------|-------------|--|
| Table 1 | N/A (no program) | | | Table was produced manually from the following sources: - International Energy Agency, 2023: "Fossil Fuel Subsidies Database." https://www.iea.org/data-and-statistics/data-product/fossil-fuel-subsidies-database . - World Bank, 2023: "World Development Indicators." World Bank, Washington, DC. https://databank.worldbank.org/source/world-development-indicators . - International Monetary Fund). 2023. "Fossil Fuel Subsidies Database." https://www.imf.org/-/media/Files/Topics/energysubsidies/EXTERNALfuelsubsidiestemplate2023new.ashx |
| Table 2 | N/A (partial program, "0 - GMD input.do") | | | Source: ITU 2023, Household surveys in Datalibweb Datalibweb dependencies are not included in the reproducibility package, the code does not reproduce this exhibit |
| Table 3 | "2 - 2312_Table3_FigureC8_TableC1_v1.do" | 243 | Table3.tex | |
| Table 4 | "2 - 2312_Body_v1.do" | 220 | Table4.tex | |
| Table 5 | "2 - 2312_Body_v1.do" | 329 | Table5.tex | |
| Figure 1 | N/A (no program) | | | Figure from authors' survey |
| Figure 2 | "3 - basic stat WP paper.do" | 394 | Figure2.png | |

| | | | | |
|-----------|---|-----|--|------------------------|
| Figure 3 | "3 - basic stat WP paper.do" | 403 | Figure3.png | |
| Figure 4 | "3 - basic stat WP paper.do" | 427 | Figure4.png | |
| Figure 5 | "2 - 2312_Body_v1.do" | 585 | Figure5.png | |
| Figure 6 | "3 - basic stat WP paper.do" | 459 | Figure6.png | |
| Figure 7 | "3 - basic stat WP paper.do" | 470 | Figure7data.xlsx | Graph is done in Excel |
| Figure 8 | "3 - basic stat WP paper.do" | 476 | Figure8.png | |
| Figure 9 | "4 - 3. Analysis.do" and "Wordclouds.R" | 690 | Figure9_keywords_q26b_(Fullsample).png | |
| Figure 10 | "4 - 3. Analysis.do" | 831 | Figure10.png | |
| Figure 11 | "2 - 2312_Body_v1.do" | 606 | Figure11a.png, Figure11b.png | |
| Figure 12 | "2 - 2312_Body_v1.do" | 754 | Figure12a.png, Figure12b.png | |
| Figure 13 | "2 - 2312_Body_v1.do" | 905 | Figure13.png | |
| Figure 14 | "2 - 2312_Body_v1.do" | 944 | Figure14.png | |

Appendix A

| Figure/Table # | Program | Line number | Output file | Note |
|----------------|------------------------------------|-------------|--------------|-----------------------------|
| Table A1 | "3 - basic stat WP paper.do" | 378 | TableA1.tex | |
| Table A2 | "3 - basic stat WP paper.do" | 360 | TableA2.tex | |
| Figure A1 | N/A (no program) | | | Figure from authors' work |
| Figure A2 | N/A (no program) | | | Figure from authors' survey |
| Figure A3 | "2 - 2312_Appendix (tables)_v1.do" | 1174 | FigureA3.png | |

| | | | | |
|-----------|------------------------------|-----|--------------|--|
| Figure A4 | "3 - basic stat WP paper.do" | 497 | FigureA4.png | |
|-----------|------------------------------|-----|--------------|--|

Appendix C

| Figure/Table # | Program | Line number | Output file | Note |
|----------------|--|-------------|---|---------------------------------------|
| Table C1 | "2 - 2312_Table3_FigureC8_TableC1_v1.do" | 288 | TableC1.xlsx | Need correct version of "iebalta" ado |
| Table C2 | "2 - 2312_Appendix (tables)_v1.do" | 224 | TableC2.tex | |
| Table C3 | "2 - 2312_Appendix (tables)_v1.do" | 333 | TableC3.tex | |
| Table C4 | "2 - 2312_Appendix (tables)_v1.do" | 562 | TableC4.tex | |
| Table C5 | "2 - 2312_Appendix (tables)_v1.do" | 627 | TableC5.tex | |
| Table C6 | "2 - 2312_Appendix (tables)_v1.do" | 752 | TableC6.tex | |
| Table C7 | "2 - 2312_Appendix (tables)_v1.do" | 817 | TableC7.tex | |
| Table C8 | "2 - 2312_Appendix (tables)_v1.do" | 942 | TableC8.tex | |
| Table C9 | "2 - 2312_Appendix (tables)_v1.do" | 1007 | TableC9.tex | |
| Table C10 | "2 - 2312_Appendix (tables)_v1.do" | 1132 | TableC10.tex | |
| Figure C1 | "5 - Pre-analysis_c1_c2_c3.do" | 110 | FigureC1.png | Source: IEA 2023, WDI 2023 |
| Figure C2 | "5 - Pre-analysis_c1_c2_c3.do" | 157 | FigureC2.png | Household surveys from Datalibweb |
| Figure C3 | "5 - Pre-analysis_c1_c2_c3.do" | 125 | FigureC3.png | IMF 2021 |
| Figure C4 | "3 - basic stat WP paper.do" | 508 | FigureC4.png | |
| Figure C5 | "3 - basic stat WP paper.do" | 525 | FigureC5.png | |
| Figure C6 | "3 - basic stat WP paper.do" | 541 | FigureC6.png | |
| Figure C7 | "2 - 2312_Appendix (tables)_v1.do" | 1187 | FigureC7.png | |
| Figure C8 | "2 - 2312_Table3_FigureC8_TableC1_v1.do" | 275 | FigureC8a.png, FigureC8b.png, FigureC8c.png | |
| Figure C9 | "2 - 2312_Appendix (tables)_v1.do" | 1232 | FigureC9.png | |

Appendix D

| Figure/Table # | Program | Line number | Output file | Note |
|----------------|---|-------------|---|---------------------------|
| Table D1 | "4 - 3. Analysis.do" | 524 | TableD1.xlsx | |
| Table D2 | "4 - 3. Analysis.do" | 582 | TableD2.xlsx | |
| Figure D1 | N/A (no program) | | | Figure from authors' work |
| Figure D2 | "4 - 3. Analysis.do" and "Wordclouds.R" | 663 | FigureD2_data_themes_q26b (Full).png | |
| Figure D3 | "4 - 3. Analysis.do" and "Wordclouds.R" | 702 | FigureD3_data_keywords_q31 (Full).png | |
| Figure D4 | "4 - 3. Analysis.do" and "Wordclouds.R" | 666 | FigureD4_data_themes_q31 (Full).png | |
| Figure D5 | "4 - 3. Analysis.do" | 778 | FigureD5.png | |
| Figure D6 | "4 - 3. Analysis.do" | 792 | FigureD6.png | |
| Figure D7 | "4 - 3. Analysis.do" | 980 | FigureD7.png | |
| Figure D8 | "4 - 3. Analysis.do" | 1020 | FigureD8.png | |
| Figure D9 | "4 - 3. Analysis.do" | 1059 | FigureD9.png | |
| Figure D10 | "4 - 3. Analysis.do" | 929 | FigureD10_1.png, FigureD10_2.png, FigureD10_3.png, FigureD10_4.png | |

