Readme File for Replication Package for Mexico Poverty Assessment Chapter 1 and Annexes

Overview

This document provides information needed to replicate the results in "Mexico Poverty Assessment Chapter 1 and Annexes 1, 2 and 3."

The folders included in this replication package are the following:

- ado
- do
    - Panel ENOE do
- dta
    - raw
- log
- output
    - Figures and Tables in Excel

All of them should be in the same folder, "Reproducibility Package Chapter 1 PA (Poverty and Equity Assessment) Mexico" to run the project do files below.

Computational requirements

Software Requirements: Stata version 14 or higher is required. Packages for the analysis are downloaded from the website and saved in the ado files. The package used is DASP Version 3.03 (Distributive Analysis Stata Package by Abdelkrim Araar and Jean-Yves Duclos (2022).

Memory and Runtime Requirements

The analyses can be run on a standard desktop computer, except for the "Panel ENOE do" inside the folder. The folder contains data, do files, and a readme that explains the requirements and the steps to generate the panel. The runtime is between 5 and 6 hours.

Data Source:

The data used in this chapter and the annexes are sourced from various publicly available sources, including the World Development Indicators (WDI), Poverty and Inequality Platform (PIP), Socio-Economic Database for Latin America and the Caribbean (SEDLAC), ILO, OECD, CONEVAL, among others. Data not sourced from SEDLAC is available in the folder "Figures and Tables in Excel" folder within the "output" folder.

Database Name	**File Name in Package**	**Source**	**Link**
National Survey of Household Income and Expenditure (ENIGH)	dta/raw/years/	Consulted through datalib web. See section datalib web below.	https://en.www.inegi.org.mx/programas/enigh/nc/2022/

National Survey of Occupation and Employment (ENOE)	MEX_2005_2023_ENOE_V01_M_V06_A_GLD_FULLSAMPLE.dta; MEX_2005_2023_PANEL_QUARTER.dta	Do files to create it are available at do/Panel ENOE do/, but intermediate data is included due to large memory requirements.	https://en.www.inegi.org.mx/programas/enoe/15ymas/#microdata
INEGI	CPI.xlsx	INEGI	https://www.inegi.org.mx/temas/inpc/#tabulados
World Bank Data (PPP Conversion Factors)	wb_ppp_conversion_factors.dta	World Bank	https://data.worldbank.org/indicator/PA.NUS.PPP?locations=MX
Coneval	pobreza_22.dta	Coneval, Bases de datos y programas de cálculo 2022 (STATA)	https://www.coneval.org.mx/Medicion/MP/Paginas/Programas_BD_2022.aspx
INEGI	geo.dta	Border zone Municipalities	https://www.inegi.org.mx/app/ageeml/

To replicate the results presented in the report, follow the steps outlined below:

STEP 1: Ensure the availability of Datasets

Ensure that you have access to "datalibweb". This will let you download the dataset and do most of the analysis. If you do not have access to it, please check the instructions in *Data Code Report Availability Statement Mexico Chapter 1 document*.

The raw data comes from SEDLAC in dta format, and it will be saved in "raw" folder, within "dta" folder. The datasets are from the year 2000 through 2022. In addition, there are two more datasets in the "dta" folder: i) geo.dta, and ii) pobreza_22.dta that are used to analyze the Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH).

For the analysis using ENOE, the data set are i) MEX_2005_2023_ENOE_V01_M_V06_A_GLD_FULLSAMPLE.dta, ii) MEX_2005_2023_PANEL_QUARTER.dta, iii) wb_ppp_conversion_factors.dta and iv) CPI.xlsx. The details of how the first two datasets are constructed can be found in the folder "do/Panel ENOE do", where the Readme_PA Panel ENOE explains step by step and provides the do file needed.

STEP 2: Set-up the folders directory and root file

i) Open the do file master_PA Mexico Chapter 1.do
ii) Set up the root file directory on line 12
iii) It will run the necessary ado files

STEP 3: Execute master_PA Mexico Chapter 1.do file

Data is downloaded from datalibweb. Please read "Data Availability Statement PA Mexico Chapter 1", running do file "download data.do" This do file will download the raw data and save it in the raw folder, inside the dta folder for later use.

Data is cleaned and prepared for analysis by running the do file "0 Poverty Data.do". This do file will generate the "ENIGH_'*year*'_to_use.dta" for the following years: 2022, 2020, 2018, 2016, 2014, 2008, 2006, and 2000, which will be used in the analysis.

The master do file will subsequently call the do-files and generate the figure or table in order.

Figures can be found in the Excel **FINAL figures For MEX PEA.xlsx**

- Figure 1.
  - Figure 1 (left panel), Figure 1 (right panel). Data from sheet *Figure 1 data*. This data is generated by running the do file "1.0 Poverty Mexico - PovTrends.do" using data processed by SEDLAC.
- Figure 2.
  - Data from sheet *Figure 2 data*, and *Figure 2 data (2)*. This data is generated by running the do file "2 Poverty Mexico - Vulnerability and MC PovTrends.do" using data processed by SEDLAC. The poverty rate data comes from sheet *Figure 1 data*.
- Figure 3
  - Figure 3 (top right), Figure 3 (top right), Figure 3 (bottom left) and Figure 3 (bottom right) data comes from sheet *data fig 3 & 5 & 7.* This data is generated by running the do file "4 Poverty in Mexico-HuppiRavallion.do". Results are shown in the log file Huppi-Ravallion.smcl in the folder "log". The numbers in the graphs are the population numbers for the "Poverty Component" and "Population Component" of the decomposition output table "Population shades and FGT indices" from Stata. An example of how to read the Stata output is in the same excel.
- Figure 4
  - Figure 4 (top right), Figure 4 (top left), Figure 4 (bottom left), and Figure 4 (bottom right) data come from sheet *data fig 4*. This data is generated by running the do file "3 Poverty-Mexico-growth incidence curves.do"
- Figure 5
  - Data from sheet *data fig 3 & 5 & 7*. This data is generated by running the do file "5 1 Poverty-Mexico – PaesDeBarros.do". The results are exported to Excel in the table "Changes in Poverty by Income Source" in cell H55.
- Figure 6 (excel)
  - Figure 6 (left), Figure 6 (center) and Figure 6 (right). Data from *sheet data fig 6*. Data is downloaded from INEGI, Encuesta Nacional de Ocupacion y Empleo using this link, and data on the ratio of minimum wage to median wage from the OECD Data Explorer using this link. The original data can be found in the folder "Figures and Tables in Excel", inside the folder output in the Excel "Original inputs for table 6 (from INEGI and OECD).xlsx

- Figure 7

- Figure 7 (left) Data from sheet “data fig 3 & 5 & 7”. This data is generated by running the do file “5 2 Poverty-Mexico – PaesDeBarros sex.do" The results are exported to excel in the table “Changes in Poverty by Income Source: Gender” in the cell H72.
- Figure 7 (center) Data from sheet “data fig 3 & 5 & 7”. This data is generated by running the do file “5 3 Poverty-Mexico – PaesDeBarros informal.do" The results are exported to excel in the table “Changes in Poverty by Income Source: Formal” in the cell H90.
- Figure 7 (right) Data from sheet “data fig 3 & 5 & 7”. This data is generated by running the do file “5 4 Poverty-Mexico – PaesDeBarros MW.do" The results are exported to excel in the table “Changes in Poverty by Income Source: MW” in the cell H108. Minimum wages are extracted from Comision Nacional de los Salarios Mínimos CONSAMI (https://www.gob.mx/conasami/documentos/tabla-de-salarios-minimos-generales-y-profesionales-por-areas-geograficas)

- Figure 8
  - Figure 8 (left) Data from the sheet “data fig_8”. This data is generated by running the do file “6 Chronic and transient labor poverty.do”
  - Figure 8 (right) Data from the sheet “data fig_8”. This data is generated by running the do file “6 1 Chronic poverty a la Ravallion for ENOE panels.do”
- Figure 9
  - Figure 9 (right) Data from the sheet “data fig 9 (right)”. This data is generated by running the do file “7 Inequality in Mexico - InequalityTrends.do"
  - Figure 9 (left) Data from the sheet “data fig 9”. This data is generated by running the do file “8 Inequality in Mexico - InequalityTrends-95 .do"
- Figure 10
  - Figure 10 (top left), Figure 10 (bottom left), Figure 10 (top right) Figure 10 (bottom right) Data from the sheet “data fig 10 & 11”. This data is generated by running the do file “6 Inequality in Mexico - Mookherjee-Shorrock.do”
- Figure 11
  - Figure 11. Data from the sheet “data fig 10 & 11”. This data is generated by running the do file “10 Inequality Mexico - PaesDeBarros.do”
- Figure 12 (excel)
  - Figure 12 (left) Data from the sheet “data Fig 12 (left)” This data comes from World Development Indicators using this link, using variable “GDP per capita, PPP (constant 2017 international $)”
  - Figure 12 (right) Data from the sheet “data Fig 12 (right)”. This data comes from World Development Indicators using this link, using variable “GDP per capita, PPP (constant 2017 international $)”
  - The original data can be found in the folder “Figures and Tables in Excel”, inside the folder output in the Excel “Original inputs for table 12 (from WDI).xlsx “

- Figure 13 (excel)
  - Figure 13 (left) Data from the sheet data Fig 13 (left). This data comes from the World Development Indicator using this link and extracting the variable “GDP per capita, PPP (constant 2017 international $) and from Poverty and Inequality Platform using this link

and extracting the variable “Poverty headcount ratio at $6.85 a day (2017 PPP) (% of population) “

- Figure 13 (right) Data from the sheet “data Fig 13 (right). This data comes from the World Development Indicator using this link and extracting the variable “GDP per capita, PPP (constant 2017 international $) and from Poverty and Inequality Platform using this link and extracting the variable “Gini coefficient. “
- The original data can be found in the folder “Figures and Tables in Excel”, inside the folder output in the Excel “Original inputs for table 13 (from WDI and PIP).xlsx “

- Figure 14 (excel)
  - Figure 14 (top left) Data from sheet “data Fig 14 (top left)”. This data comes from the World Development Indicator using this link and extracting the variable “GDP per capita, PPP (constant 2017 international $) and the OECD Data Explorer using this link and extracting the variable minimum relative to median wages of full-time workers.
  - Figure 14 (bottom left) Data from sheet “data Fig 14 (bottom left)” This data comes from the ILO-Social Security Inquiry Database using the variable: Country, Total, Work Injury: “Contingency: Employed covered in the event of work injury” for years 2021 and 2019.
  - Figure 14 (top right) Data from sheet “data Fig 14 (top right)” This data comes from the World Development Indicator using this link and extracting variable “Employment to population ratio, 15+, female (%) (national estimate)”
  - Figure 14 (bottom right) Data from sheet “data Fig 14 (bottom right)” This data comes from the World Development Indicator using this link and extracting the variable “Employment in agriculture (% of total employment) (modeled ILO estimate)” and the variable “Agriculture, forestry, and fishing, value added per worker (constant 2015 US$)”
  - The original data can be found in the folder “Figures and Tables in Excel”, inside the folder output in the Excel “Original inputs for table 14 (from WDI, ILO and OECD).xlsx “

- Annex Figure 1. The figures are generated running the do file “Annex FGT Curves 2022.do” It generates two graphs that are saved in the folder output: “Anexo_FGT_Diff_ppp17.gph” (right) and “Anexo_FGT_ppp17.gph” (left)
- Annex Figure 2. The figures are generated running the do file “Annex FGT Curves scale equivalence 2022.do" It generates two graphs that are saved in the folder output: “Anexo_FGT_Diff_ppp17_se.gph” (right) and “Anexo_FGT_ppp17_se.gph” (left)
- Annex Figure 3. The figures are generated running the do file “Annex FGT Curves without imputed rent 2022.do" It generates two graphs that are saved in the folder output: “Anexo_FGT_Diff_ppp17_woir.gph” (right) and “Anexo_FGT_ppp17_woir.gph” (left)
- Annex Figure 4. Data from the sheet “data A-fig 4”. This data comes from the CONEVAL using these links for old series (pdf) and new series.
  - The original data can be found in the folder “Figures and Tables in Excel”, inside the folder output in the Excel “Original inputs for Annex figures 4 & 5 (from WDI, OECD and CONEVAL).xlsx “
- Annex Figure 5. Data from the “data A-fig 5” This data comes from the CONEVAL using this link (pdf) and from the World Development Indicator using this link , extracting Consumer price index

(2010 = 100), PPP conversion factor, GDP (LCU per international $), PPP conversion factor, private consumption (LCU per international $) from OECD using this link.

- The original data can be found in the folder "Figures and Tables in Excel", inside the folder output in the Excel "Original inputs for Annex figures 4 & 5 (from WDI, OECD and CONEVAL).xlsx "

Tables can be found in the excel **FINAL tables For MEX PEA.xlsx**

- Table 1. This data is generated by running the do file "11 Poverty in Mexico -DattRavallion.do"
- Table 6 Annex. Data from sheet "Annex table 6". This data is generated by running do file "Annex Table 6 Demographic characteristics of selected ENOE panel and comparison full cross section.do"
- Table 7 Annex Data from sheet "Annex table 7" This data is generated by running do file "Annex Table 7 Poverty Vulnerability and Middle Class profiles.do"
- Table 8 Annex. Data from sheet "Annex table 7". This data is generated by running do file "Annex Table 8 Demographic profile of chronic poor, transitory poor and no poor.do"

Data Availability Statement

The data used in the study are publicly available in the reproducibility package, except for the household surveys microdata from the Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank), which are only available to members of the World Bank Group.

Datalibweb access to data

DATALIBWEB is a data system specifically designed to enable users to access the most up-to-date versions of non-harmonized (original/raw) and harmonized datasets of different collections across Global Practices. It can perform computations relevant to poverty and shared prosperity analysis based on the microdata from various harmonized collections: EAPPOV, ECAPOV, MNAPOV, SARMD, SEDLAC, LABLAC, SSAPOV, and the global collection GPWG/GMD. It also helps users have access to the original/raw data for further country-specific analysis.

There is both the STATA program <datalibweb> as well as the website with the furl "datalibweb/" for subscribing.

Each person who wants to download data in Stata using the program code <datalibweb> needs to subscribe personally within the "datalibweb/" website and sign the General Disclaimer on Terms of Use (see below).

If you have not installed DATALIBWEB, please follow these instructions:

1. type in the intranet "datalibweb/",
2. go to the tab "about",
3. follow the instructions below "manual installation":

   a. the "automatic" option is not working, so please follow the instructions related to "manual installation."

   b. follow the steps exactly. If you change the path or modify the instructions, <datalibweb> may not work.

c. restart your pc after manual installation of the package,

d. if you have problems related to the installation, please email the <datalibweb> team at: datalibweb@worldbank.org

After <datalibweb> is installed, you must:

1. accept the General Disclaimer in the datalibweb/ website (the red box in the upper right corner needs to turn green)
2. check if the data you want to access is public or private,

a. if public: you have access immediately via STATA,

b. if private: you must request access, with an explanation for the use of the data, and the TTL will have to approve (the TTL has the MOU with the country NSO in these cases),

3. please note you can search for SEDLAC or LABLAC data under "Select Server/Country",

In STATA:

1. Once you receive access, you can then use <datalibweb> within STATA to call the microdata.

2. Within STATA, type "help datalibweb" to see examples of the syntax/program code that loads the microdata into STATA.

a. Example: datalibweb, country(mex) years(2022) type(SEDLAC-03) mod(all) clear

b. This code opens the "all" module for Mexico 2022 from version 03 of the SEDLAC project.

Contact: For further queries, please reach out to Mariel Cecilia Siravegna (msiravegna@worldbank.org) or to Samuel Freije-Rodriguez (sfreijerodriguez@worldbank.org)