

Replication material for “Raising Children Alone in Latin America and the Caribbean: Strong Mothers and Weak Social Assistance”

Laura Cuesta , Jacobus de Hoop , Hugo Ñopo

This folder contains all Stata codes and datasets necessary to replicate the results of this paper. The empirical analysis in this paper relies primarily on SEDLAC household surveys datasets for Latin America and the Caribbean countries, which are only available to members of the World Bank Group.

Overview

This document provides information needed to replicate the results in the paper. The folder structure is as follows:

- *code*: Holds all the Stata .do files.
- *input*: Holds cleaned dataset for running the analysis. Additionally, it includes a couple of raw datasets that are not available through Datalibweb because of their size.
- *output*: Holds all the excel files with the results to produce tables and figures, figures produced by Stata and an excel file with the final version of the tables and figures.

1- Description of data

• Data sources

One data source was used to construct the analytics for this paper, that is SEDLAC¹.

SEDLAC database includes household surveys microdatasets from all Latin American and some Caribbean countries, which are only available to members of the World Bank Group through [Datalibweb portal](#). Datalibweb package should be installed in Stata in order to access this data (please refer to this user guide: [Datalibweb user guide – Stata interface](#)). Accessed date is January 2025.

Note: some datasets require submitting an access request, namely for Dominican and Honduras. Additionally, some others are not available online because of their size (Colombia, 2021 and 2022)

Information taken from this database should be cited as “Source: SEDLAC (CEDLAS and The World Bank)” or “Source: Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank).”

¹ Socio-Economic Database for Latin America and the Caribbean (CEDLAS and The World Bank)

- **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 for 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 can help users to 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 that 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). To assure that users have the most up to date microdata, and to make sure outputs/projects can be replicated by the original user or others, we highly recommend that people do not download data onto their computers, but rather form the habit of accessing it using the <datalibweb> program in STATA (incorporating it into their do-file codes), as explained further below. Users should always note the vintage of the data.

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(ecu) years(2017) type(SEDLAC-03) mod(pov) clear
 - b.This code opens the POVerty module for Ecuador 2017 from the version 03 of the SEDLAC project.

• Summary of Data

The SEDLAC microdatasets are called through Datalibweb and downloaded to the local hard drive, in the folder named “input”. There are two subfolders: 2022 and historical. Using those inputs, some microdatasets for Latin America and the Caribben are created, which include the necessary data to produce the tables and figures.

File	Source	Data	Notes
<i>input/2022/lac_circa_2022.dta</i>	Constructed from SEDLAC	Yes	<i>01_download data (2022).do-</i> .do file to download datasets <i>02_creation of dataset (2022).do-</i> .do file to create the aggregated dataset
<i>input/2022/lac_circa_2022_variables.dta</i>	Constructed from “lac_circa_2022.dta”	Yes	<i>03_variables for tables.do</i> - .do file for construction Dataset used to produce tables
<i>input/historical/short_circa_`year'.dta</i>	Constructed from SEDLAC	Yes	<i>01_download data (historical_trends).do-</i> .do file to download datasets and create the aggregated ones
<i>input/ historical/short_circas.dta</i>	Constructed by pooling <i>short_circa_`year'.dta</i> datasets for all the years	Yes	<i>02_creation of dataset (historial_trends).do-</i> .do file for pooling Dataset used to produce figures

2- Instructions for replicators

All the .do files can be run by the *00_master.do*. You only need to change the path directory at the top of the *00_master.do* file. Although the package includes two do-files to download the microdata for each country, the package also includes those dataset, which can be found in the

folder named “input”. That is the reason why the lines 50 and 51 of *00_master.do* are “muted”. The 2 clean datasets used to produce the output tables and figures are *lac_circa_2022_variables.dta* and *short_circas.dta*. Other numbers mentioned in the paper are constructed using *04_other numbers.do*

Note 1: The final formatting of the Tables and the construction of figures 1-3 is done in Excel using the exported results. Some small changes in the order of some columns/rows of exported results are required. The calculated stars in tables 2-6 are based on calculated p-values included in each spreadsheet. A p-value ≤ 0.01 responds to ***, a $0.01 < \text{p-value} \leq 0.05$ responds to **, a $0.05 < \text{p-value} \leq 0.10$ responds to *, and p-value > 0.10 responds to no star.

Note 2: We have provided the datasets for Colombia 2021-2022, because they cannot be called through Datalibweb. They can be found in the folder named “input”.

Note 3: In case of experiencing issues in internet connection/other errors that interrupted your data call from Datalibweb (*01_download data (2022).do* & *01_download data (historical_trends).do*), you must run the loop again on an empty dataset and not on the old, interrupted dataset. Otherwise, the data will be messed up completely. Don’t use “*drop duplicates, force*”. Always start on clean dataset even if it means running the data loop from scratch.

3- List of Exhibit

Exhibit name	Output filename	Script	Note
Table 1	table1.xls	03_table_1.do	
Table 2	table2.xlsx	03_table_2.do	
Table 3	table3.xlsx	03_table_3.do	
Table 4	table4.xlsx	03_table_4.do	
Table 5	table5.xlsx	03_table_5.do	
Table 6	table3.xlsx	03_table_6.do	
Table A1	tableA1.xlsx	03_table_A1.do	Final version can be found in output/Mothers who raise children alone - tables and charts (FINAL).xlsx
Table A2	tableA2.xlsx	03_table_A2.do	
Figure 1	figure 1.xlsx	03_figures.do (line 226-263)	
Figure 2	figure 2.xlsx	03_figures.do (line 265-269)	
Figure 3	figure 3.xlsx	03_figures.do (line 271-275)	
Figure 4	figure 4a.gph & figure 4b.gph	03_figures.do (line 284-315)	

4- Requirements

• Software requirements

Stata is used for the analysis in this project. This project is written in Stata 19.5 and has not been tested on older versions.

The guidelines below, sourced from <https://www.stata.com/products/compatible-operating-systems/>, outline the following computational requirements:

Platforms

Stata for Windows®

- Windows 11*
- Windows 10 *
- Windows Server 2022, 2019, 2016, 2012R2 *

* Stata requires 64-bit Windows for x86-64 processors made by Intel® or AMD (Core i3 equivalent or better)

[Find out if your OS is 64-bit compliant.](#)

Stata for Mac®

- Mac with Apple Silicon or Intel processors
- macOS 11 (Big Sur) through macOS 26 (Tahoe) for Macs with Apple Silicon and macOS 10.13 (High Sierra) through macOS 26 (Tahoe) for Macs with Intel processors

Stata for Linux

- Any 64-bit (Core i3 equivalent or better) running Linux
- Minimum requirements include the GNU C library (glibc) 2.28 or better and libcurl4
 - Check the output of `ldd -v` within a terminal
- For xstata, you need to have GTK 2.24 installed

Hardware requirements

Package	Memory	Disk space
Stata/MP	4 GB	4 GB
Stata/SE	2 GB	4 GB
Stata/BE	1 GB	4 GB

Stata for Linux requires a video card that can display thousands of colors or more (16-bit or 24-bit color)

• Packages dependencies

None

• Memory and Runtime and Storage Requirements

It takes around one hour and a half to run the whole package, excluding the do-files to download the data (*01_download data (2022).do* & *01_download data (historical_trends).do*). It takes around 20 minutes to run the analysis codes (those named “*03_figures*”, “*03_table_#*” and “*04_other_numbers*”).

The final version of the datasets for the analysis (i.e. *short_circas.dta* and *lac_circa_2022_variables.dta*) are 3.6 GB in size.

5- Code description

The code comprises 17 do-files, organized in five different stages.

The zero stage (00) is the master do-file. It sets the directories, list of countries and main variables. This file also calls the rest of the do-files in the correct order.

The first stage (01) calls the datasets through Datalibweb, except in the case of those for Colombia 2021 and 2022, and save them in the local hard drive.

The second stage (02) creates the two datasets for Latin America and the Caribbean, which are needed for the analysis.

The third stage (03) creates the main variables of the analysis in the case of the dataset for 2022.

The fourth stage (04) produces the output data to create the tables and figures included in paper. It should be mentioned that some variables are created in this file, because they are specific to the analysis performed in each do-file.

The fifth stage (05) computes additional results included in the paper, but not in a table or figure.