

# README: Replication Files for "*Breadwinners and Caregivers: Examining the Global Relationship between Gender Norms and Economic Behavior*"

## Overview

This repository contains the necessary files for replicating the analysis in "*Breadwinners and Caregivers: Examining the Global Relationship between Gender Norms and Economic Behavior*" by Markus Goldstein, Paula Gonzalez, Talip Kilic, Sreelakshmi Papineni and Philip Wollburg.

*Memory and Runtime Requirements:* The Stata analysis code approximately requires 18 minutes to execute completely.

*Software Requirements:* Stata (Tested on version 18)

Required Stata packages:

- estout (3.23)
- iefieldkit (2.0)
- ietoolkit (6.3)
- coefplot (1.8.3)
- spmap
- shp2dta

## Instructions for Replicators

1. **Prepare Environment:** Open the master dofile (MasterDofile\_BreadWinnersCaregivers.do). Ensure user settings, including unique user number and OneDrive paths, are correctly configured.
2. **Install Packages:** If necessary, install user-written packages by modifying line 37 in the master dofile (change 0 to 1). This is a one-time setup per computer.
3. **Set Repository Path:** Modify line 64 to reflect the path of the repository on your computer.
4. **Execute Analysis:** The master dofile (MasterDofile\_BreadWinnersCaregivers.do) is preconfigured for immediate execution. However, for a full analysis, you may need to set/check the flags (DataConstruction, DataAnalysis) to 1. After adjusting these flags, run the master dofile. For running individual sections, simply set the respective flags for those specific sections.
5. **Generate Outputs:** Running the master dofile will construct the main datasets for analysis and generate all tables and graphs from the working paper.
6. **Output Location:** Outputs will be stored in \Outputs\Raw\Figures and \Outputs\Raw\Tables. Main analysis outputs are in the 'Main' folder, and appendix outputs in the 'Appendix' folder. All outputs are compiled in two main files containing all tables.

## Datasets Description

**Accessing the data:** The paper utilizes data from a collaboration with Facebook Data for Good and therefore requires prior approval from Meta to access the individual-level dataset. This can be done via the World Bank's Data Development Partnership (DDP) license agreement granted on a project-by-project basis. For

use of the data, colleagues within the World Bank need to submit a request through the Partnership portal – doing so sends the request directly to Meta, which can then approve the new data use case (even if it is just reproducing what you did before). With Meta's approval, the DDP team can then provision the data to the Bank staff. Since the license only covers WB staff requests, no other external parties could be granted access to the data. For Bank staff access to the data, advise to submit a request here: <https://portal.datapartnership.org/submitproposal>

**Principal Dataset:** The "*Gender At Home Raw Dataset*" is a comprehensive dataset at the individual level, consisting of 461,748 observations. Due to data sharing restrictions, only a cleaned and anonymized subset of this dataset is available for sharing.

The original dataset should be located at "\$data\Raw\GenderAtHomeDataRaw\_092021.dta". This is the dataset in stata format received from the Development Data partnership data\_v3\_sogeh\_microdata\_2020.

This subset includes data from respondents who reported their gender, representing approximately 50% of the original sample. The dataset can be accessed at \$data/Intermediate/*GenderAtHome\_DataClean*. This forms the starting point for all subsequent analyses.

Following the execution of the master data construction dofile, several derived datasets are generated for use in various analyses throughout the paper:

1. \$data/Final/*GenderAtHome\_DataAnalysisInd.dta*: This is the primary dataset for individual-level analysis.
2. \$data/Final/*GenderAtHome\_DataAnalysisCountry.dta*: This dataset aggregates key variables at the country level and is primarily used for generating graphs.
3. \$data/Final/*GenderAtHome\_DataAnalysisMaps.dta*: Created specifically for reproducing the maps presented in the paper.

**Secondary Datasets:** Located in data/Intermediate/OtherDataset, these datasets include a range of publicly available administrative and individual data sources. These secondary datasets have been utilized for a more comprehensive analysis in the study. The original sources of these datasets are provided, enabling replicators to access them directly.

1. WVS\_Longitudinal\_1981\_2020: <https://www.worldvaluessurvey.org/WVSDocumentationWVL.jsp>
2. DFID2020\_DataClean
3. Gallup Data: Data can be found in the appendix A2 to A14 of the paper "How are gender norms perceived? (Bursztyn, et al 2023). Link: [https://www.nber.org/system/files/working\\_papers/w31049/w31049.pdf](https://www.nber.org/system/files/working_papers/w31049/w31049.pdf).
4. WBI Data: We download the data for GDP and labor force participation at the country level for all countries in the world for 2019 and 2020: <https://databank.worldbank.org/source/world-development-indicators>

The database is organized in a long format, where each row represents a country, and each column contains various indicators. The indicators considered are as follows:

Variable Name	Variable Label
countrycode	Country Code
year	Year

se_adt_1524_lt_fe_zs	Literacy rate, youth female (% of females ages 15-24)
se_adt_litr_fe_zs	Literacy rate, adult female (% of females ages 15 and above)
se_prm_cuat_fe_zs	Educational attainment, at least completed primary, 25+ years, female
se_prm_tenr_fe	Adjusted net enrollment rate, primary, female (% of primary school age children)
se_prm_tenr_ma	Adjusted net enrollment rate, primary, male (% of primary school age children)
se_sec_cuat_lo_fe_zs	Educational attainment, at least completed lower secondary, 25+, female
se_sec_cuat_po_fe_zs	Educational attainment, at least completed post-secondary, 25+, female
se_sec_cuat_up_fe_zs	Educational attainment, at least completed upper secondary, 25+, female
sg_tim_uwrk_fe	Proportion of time spent on unpaid domestic and care work, female (% of 24 hour
sg_tim_uwrk_ma	Proportion of time spent on unpaid domestic and care work, male (% of 24 hour d
sl_tlf_cact_fe_zs	Labor force participation rate, female (% of female population ages 15+) (model
sp_dyn_tfrt_in	Fertility rate, total (births per woman)
sp_pop_totl_ma_zs	Population, male (% of total population)
fp_cpi_totl_zg	Inflation, consumer prices (annual %)
ny_gdp_mktp_kd	GDP (constant 2010 US\$)
ny_gdp_mktp_kd_zg	GDP growth (annual %)
ny_gdp_pcap_cd	GDP per capita (current US\$)
ny_gdp_pcap_kd_zg	GDP per capita growth (annual %)
rl_est	RL.EST
ge_est	GE.EST
it_net_user_zs	IT.NET.USER.ZS

- World Shape Coordinates: This item refers to the official World Bank country boundaries shapefile, which can be found by selecting "World Country Polygons - Very High Definition: wb\_countries\_admin0\_10m" from the [World Bank Data Catalog](#). The data, stored within the Data/Intermediate/OtherDatasets/world\_shape.dta file, includes the original shapefile augmented with additional variables such as map color. Additionally, the file world\_shape\_coord.dta contains the coordinates for each country, facilitating the creation of world maps by merging this data with the world shape file and any variables of interest for analysis.

## DoFile Structure and Execution Order

Master Dofile (MasterDofile\_BreadWinnersCaregivers.do): The primary script for orchestrating all other dofiles. Run this first for complete replication following steps in instructions for replicators.

### Main Sections of Dofile:

#### Part 1.

Part 1. Data Import and Cleaning (Cleaning/Part1\_MasterDataCleaning.do): While direct access to the raw dataset may not be available, we have prepared a detailed dofile for its cleaning and preparation. This dofile efficiently imports the dataset into Stata, embarking on a series of essential tasks aimed at readying the data for variable construction. These tasks cover a comprehensive scope, including adding value and variable labels, resolving special missing values (e.g., .d for "Don't know"), transforming string variables into numeric formats, and applying sample selection criteria to refine the dataset for subsequent analysis (the refined data post-selection criteria application is then available for sharing). The dofile is meticulously structured into three primary sections: data importation, general cleaning, and in-depth cleaning of individual variables. This organization ensures a systematic and efficient approach to data preparation.

Part 2. Data Variable Construction (/Intermediate/Part2\_MasterDataConstruction.do): Constructs necessary variables for analysis, merges with other datasets, creates weights, and generates country-level and map datasets.

1. Construction of Individual Variables
2. Merging with other datasets
3. Weights creation and variables related to sample selection
4. Generating dataset at the country level.
5. Generating dataset for Maps

Part 3. Data Analysis (/Analysis/Part3\_MasterDataAnalysis.do): Performs main and appendix analysis, including descriptive statistics, regression tables, and graphs.

## Outputs

- i. Outputs include processed data, tables, and graphs stored in specified output directories. (\Outputs\Raw/Tables and Raw/Figures )
- ii. Each file is initially exported individually into the Main and Appendix folders. Subsequently, these files are automatically compiled into main table files, named Outputs\Raw/GenderAtHome\_MainTables\_date and GenderAtHome\_AppendixTables\_date.
- iii. In the Outputs\Final/ folder, you will find the final versions of the files mentioned in point (ii). These final versions are the result of manual formatting enhancements. It is important to note that while the format (such as font type, cell size, etc.) has been modified for improved presentation, no changes have been made to the actual values or labels in the data. Graphs are exported in both .gph and .png formats. PnG files are uploaded to LaTeX. Some images are combined into a single image for paper presentation.
- iv. Maps are produced by us and recreated by the WBG Cartographic Unit using our clean dataset, adhering to WBG Cartographic Guidelines.

## List of Analysis Codes and Outputs

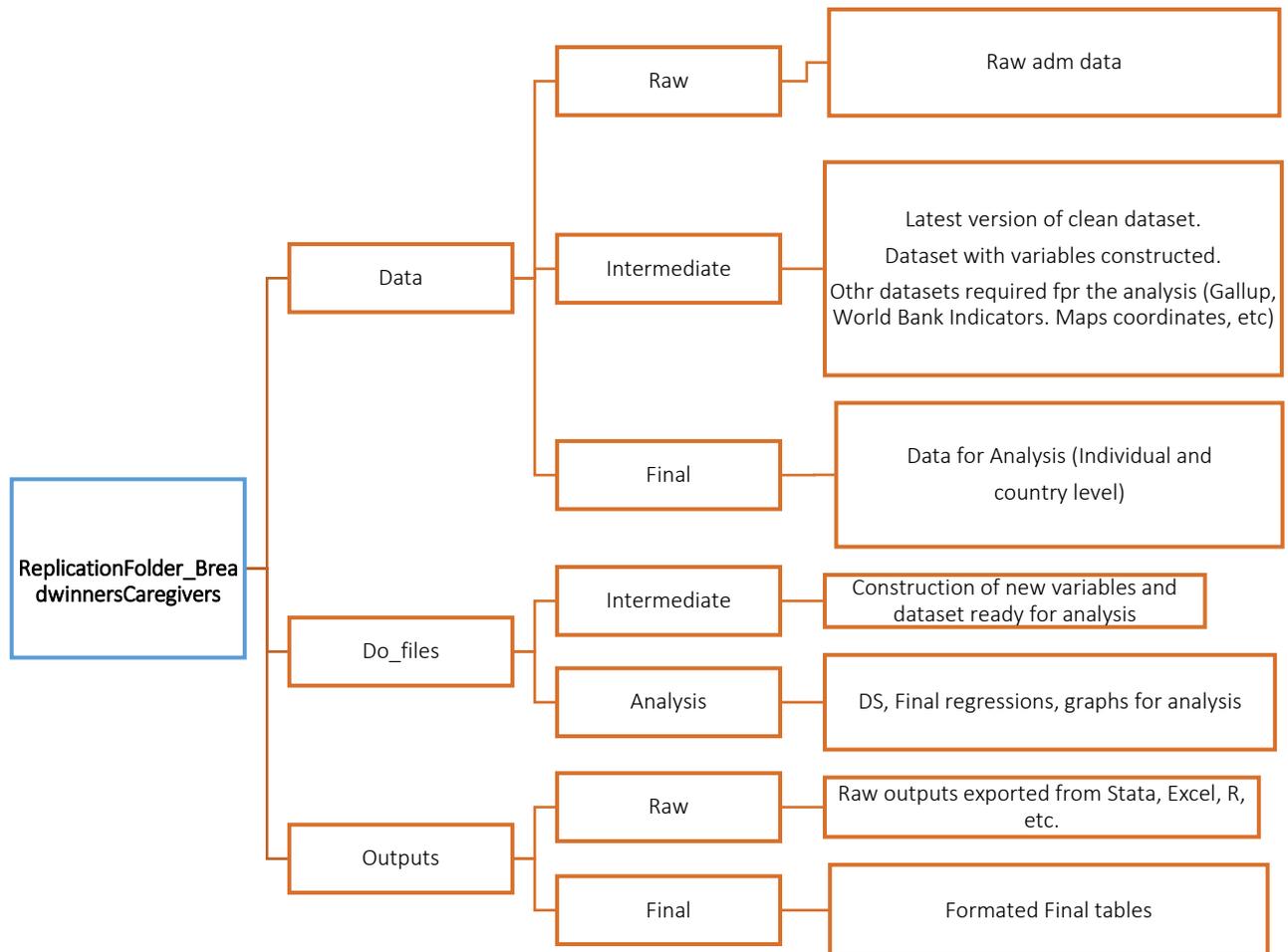
All code can be executed from the Master.do script, providing a centralized approach for running the full analysis. Additionally, codes for individual exhibits are designed to be executed independently, with the condition that folder globals and custom programs are correctly set up in both the master dofile and the respective section dofiles. This flexibility ensures the accurate reproduction of all numerical data presented in the paper. For ease of reference, a table is included which specifies the name of each output as mentioned in the paper, its location, the program responsible for its execution, and the individual output's file name in the folder.

<b>Main Analysis</b>			
Folder path for individual dofiles: \$dofiles/Analysis/Sections			
Unique file for tables: \$outputs_raw/Tables\GenderAtHome_MainTables_`svdate'.xlsx			
<b>Exhibit</b>	<b>Input dataset</b>	<b>Program</b>	<b>Individual Output file</b>
Table 1	GenderAtHome_DataAnalysisInd.dta	DS_SampleCharacteristics.do	SampleChar_Gender
Figure 1	GenderAtHome_DataAnalysisCountry.dta	Fig_GenderGapNormsRegion.do	RegionGenderNorm1/RegionGenderNorm2/RegionGenderNorm3. png
Figure 2	GenderAtHome_DataAnalysisInd.dta	Fig_NormsCountryLevel.do	norm1_part1/ norm1_part2. png
Figure 3	GenderAtHome_DataAnalysisInd.dta	Fig_NormsCountryLevel.do	Norm2_part1/ norm2_part2. png
Figure 4	GenderAtHome_DataAnalysisInd.dta	Fig_NormsCountryLevel.do	Norm3_part1/ norm3_part2. png
Figure 5	GenderAtHome_DataAnalysisMaps.dta	Fig_MapsMisperception.do	Map_norm1. png
Figure 6	GenderAtHome_DataAnalysisMaps.dta	Fig_MapsMisperception.do	Map_norm2. png
Figure 7	GenderAtHome_DataAnalysisMaps.dta	Fig_MapsMisperception.do	Map_norm3. png
Table 2a	GenderAtHome_DataAnalysisInd.dta	Regressions_LaborAndNorms.do	Labor_GenderEquality.csv
Table 2b	GenderAtHome_DataAnalysisInd.dta	Regressions_LaborAndNorms.do	Labor_MaleBreadwinner.csv
Table 2c	GenderAtHome_DataAnalysisInd.dta	Regressions_LaborAndNorms.do	Labor_FemaleCaregiver.csv
Table 3	GenderAtHome_DataAnalysisInd.dta	Regressions_LaborAndNorms.do	Labor_AllNorms.csv
Table 4a	GenderAtHome_DataAnalysisInd.dta	Regressions_DecisionMakingAndNorms	DM_GenderEquality.csv
Table 4b	GenderAtHome_DataAnalysisInd.dta	Regressions_DecisionMakingAndNorms	DM_MaleBreadwinner.csv
Table 5	GenderAtHome_DataAnalysisInd.dta	Regressions_DecisionMakingAndNorms	DM_AllNorms.csv
Table 6a	GenderAtHome_DataAnalysisInd.dta	Regressions_CareDomesticRespAndNorms.do	CareDom_GenderEquality.csv
Table 6b	GenderAtHome_DataAnalysisInd.dta	Regressions_CareDomesticRespAndNorms.do	CareDom_FemaleCaregiver.csv
Table 7	GenderAtHome_DataAnalysisInd.dta	Regressions_CareDomesticRespAndNorms.do	CareDom_AllNorms.csv

<b>Appendix Analysis</b>
Folder path for individual dofiles: \$dofiles Analysis/Sections/Appendix
Unique file compiled all tables: \$outputs_raw/Tables\GenderAtHome_AppendixTables_

<b>Exhibit</b>	<b>Input dataset</b>	<b>Program</b>	<b>Individual Output file</b>
Table A3	GenderAtHome_DataAnalysisInd.dta	DS_MainVarGender	DS_MainVarGender.csv
Table A4	GenderAtHome_DataAnalysisCountry.dta	DS_ComparissonWVS	DS_FemaleCareWVS
Table A5	GenderAtHome_DataAnalysisCountry.dta	DS_ComparissonWVS	DS_MaleBreadwinnerWVS
Table A6	GenderAtHome_DataAnalysisCountry.dta	DS_CorrBroadNormGallup	DSCorrNormGallup.csv
Table A7	GenderAtHome_DataAnalysisCountry.dta	DS_DataByRegions.do	DataEAP /DataECA.csv
Table A8	GenderAtHome_DataAnalysisCountry.dta	DS_DataByRegions.do	DataLAC/ DataMENA.csv
Table A9	GenderAtHome_DataAnalysisCountry.dta	DS_DataByRegions.do	DataNorthAmerica/DataSouthAsia/DataSSA.csv
Table A10	GenderAtHome_DataAnalysisInd.dta	DS_GenderDiffRegion.do	DS_GenderDiffRegio.ncsv
Figure B1	GenderAtHome_DataAnalysisCountry.dta	Fig_BeliefsAndNormsSplitGender.do	IndCom_GenderNorm1. png
Figure B2	GenderAtHome_DataAnalysisCountry.dta	Fig_BeliefsAndNormsSplitGender.do	IndCom_GenderNorm2. png
Figure B3	GenderAtHome_DataAnalysisCountry.dta	Fig_BeliefsAndNormsSplitGender.do	IndCom_GenderNorm3. png
Figure C1	GenderAtHome_DataAnalysisCountry.dta	Fig_GenderBeliefMacro.do	MacroNorm1. png
Figure C2	GenderAtHome_DataAnalysisCountry.dta	Fig_GenderBeliefMacro.do	MacroNorm2. png
Figure C3	GenderAtHome_DataAnalysisCountry.dta	Fig_GenderBeliefMacro.do	MacroNorm3. png
Figure D1	GenderAtHome_DataAnalysisCountry.dta	Fig_GenDiffMispNorm.do	Norm1 . png
Figure D2	GenderAtHome_DataAnalysisCountry.dta	Fig_GenDiffMispNorm.do	Norm2. png
Figure D3	GenderAtHome_DataAnalysisCountry.dta	Fig_GenDiffMispNorm.do	Norm3. png
Table D1:	GenderAtHome_DataAnalysisInd.dta	DS_DegreeMisperceptionRegion.do	DS_DegreeMispGenderRegion.csv
Table D2	GenderAtHome_DataAnalysisInd.dta	Reg_NormMisperception.do	Misp_IndividualChar.csv

## Folder Structure:



## Author Contact Information

- Author: Paula Gonzalez
- Email: pgonzalezmartine@worldbank.org