

Beyond Wages:

What Matters Most in Job Choice for Women in El Salvador

Ivette Contreras¹

Lelys Dinarte-Diaz

Amparo Palacios-Lopez

Valentina Costa

Steffanny Romero

¹Corresponding Author. The World Bank. Email: icontreras@worldbank.org.

Reproducibility Package README

Overview

The code within this replication package executes the analysis detailed in the paper "Beyond Wages: What Matters Most in Job Choice for Women in El Salvador" using Stata. This package includes all data files and codes required to go from the anonymized data to the replication of the tables included in the analysis. The master file manages the execution of all codes, anticipated to run for approximately 30 minutes. Our primary objective in this reproducibility package is to ensure both accessibility and credibility in replicating our research methodology and results.

Data Availability and Provenance Statements

To collect the relevant data for our experiment, we surveyed the sample of 1,008 households (2,480 household members) between August and October 2022. The interviews were conducted using a Computer-Assisted Personal Interviewing (CAPI) program, in which we embedded the household assignments to the treatment groups. The experiment was conducted in two regions within the Departments of San Salvador and Usulután. We selected these regions to achieve high variation in the types of activities reported. Region 1, which is located in the Department of San Salvador, includes households located close to the Metropolitan Area of San Salvador, the main economic activities of which are commerce, construction, manufacturing, and sugar cane production. Region 2 is situated in the Department of Usulután, where the main economic activities include coffee production, fishing, and commerce.

Using the 2007 Salvadoran Population Census (i.e., the most recent census available), we identified 276 enumeration areas (EAs): 114 in Region 1, and 162 in Region 2. Each of these EAs had at least 30 households and was classified as either rural or peri-urban. From the 276 EAs, we randomly selected 48 that were evenly distributed across the rural and peri-urban areas and then conducted a full household listing for each EA. We then randomly selected 21 households within each EA for a total of 1,008 households to constitute our study sample. Within each household, we randomly chose a minimum of two and up to a maximum of four working-age household members aged between 15 to 64 years, which yielded a total of 2,480 working-age individuals. The data comes from the Discrete Choice Experiment module of the survey instrument collected for the paper "Closing the Gaps: The Role of Screening Questions and Self-

Reporting in Measuring Women's and Youths' Employment and Work". We conducted a *preferences experiment* to understand the preferences and barriers related to employment attributes among women and youth. The module was administered only to self-respondents who could read and write. In total, we have data from 1,507 individuals.

Statement about Rights

- ☒ I certify that the author(s) of the manuscript have legitimate access to and permission to use the data used in this manuscript.

Summary of Availability

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

The data for this project is owned by the World Bank (it came from a survey data procured by the World Bank under the project P168212). The data will be published in the Microdata Library in July 2026 as detailed in this previous [reproducibility package](#).

Getting Started

Outline and Folder Structure

The reproducibility package consists of three folders:

- *Data anonymized*: this folder contains all the raw data files, and the final data set used for the analysis.
- *Do files*: this folder contains the related codes. The Master File labeled as “**master-dofile.do**” orchestrates the sequential execution of all code files according to the specified order.
- *Output*: this folder contains all the tables and figures reported in the paper in raw form. Output files are called appropriate names (e.g., Table2.xls) and should be easy to correlate with the manuscript.

How to Reproduce

The initial action within this reproducibility package involves accessing the master file named master-dofile.do and execute its contents. Before proceeding, ensure Stata is installed on your

system. This file establishes directory paths and executes the necessary Stata code to generate analytical outputs. Here's a concise overview of its contents:

PART 1: Install user-written packages and harmonize settings.

PART 2: Prepare folder paths and define programs.

PART 3: Run do-files.

PART 3.1: This part runs `01_combine_datasets.do` do-file that constructs variables and indicators used in the analysis.

PART 3.2: This part runs `02_main_tables.do` do-file that will generate all the main tables.

Kindly designate the user root-path in line 27 within the Master do file (`master-dofile.do`). You have the option to execute the entire do file or select specific sections. It's crucial, however, to run the do file in sequence. If opting to run only one section, ensure that preceding sections have been executed beforehand.

Dataset list

The table below lists all datasets in the raw data file required for running the do-files.

Datafile	Note	Provided
master_data.dta	Contains individual-level information from all individuals	Yes
DCE_mujeres.dta	Contains information about each image-choice set presented to the women.	Yes
section9A.dta	Contains individual-level information from DCE module	Yes

List of tables and programs

Selected tables in the paper, as explained and justified below. The regression tables produced by the Stata (raw Excel output) were subsequently manually reformatted for presentation purposes only.

Table/Figure	Program	Line Number	Output file
Table 2	Do files/ 02_main_tables.do		Table2.xlsx
Table 3	Do files/ 02_main_tables.do		Table3.xlsx
Table 4	Do files/ 02_main_tables.do		Table4.xlsx
Table 5	Do files/ 02_main_tables.do		Table5.xlsx
Table 6	Do files/ 02_main_tables.do		Table6.xlsx

Computational requirements

Software Requirements

Stata 18 should be used to run the codes. Below are the versions of each command used.

```
c:\ado\plus\m\mixlogit.ado
*! mixlogit 1.4.0 29Mar2016
*! author arh
c:\ado\plus\m\mixlogitwtp.ado
*! mixlogitwtp 1.1.0 29Mar2016
*! author arh
c:\ado\plus\o\outreg2.ado
*! outreg2 2.3.2 17aug2014 by roywada@hotmail.com
*! based on outreg 3.0.6/4.0.0 by john_gallup@alum.swarthmore.edu
```

References

Lars Vilhuber, Connolly, M., Koren, M., Llull, J., & Morrow, P. (2022). A template README for social science replication packages (v1.1). Zenodo. <https://doi.org/10.5281/zenodo.7293838>

Acknowledgment

We acknowledge the support of the International Fund for Agricultural Development and the World Bank Research Support Budget. The Institutional Review Board (IRB) of Universidad Francisco Gavidia in El Salvador reviewed and approved this research project's protocol, which is registered under ID #011-2021.