

README: Replication for “Family Matters: Social Networks and Women’s Economic Empowerment” (Nono Ayivi-Guedehoussou, Abhilasha Sahay, Maria Emilia Cucagna and Mansi Sharma, 2025)

Overview

The code in this replication package constructs all tables and figures from the working paper using Stata. The package consists of three Stata do-files, and one raw dataset is included. Replication is handled via a single master script (“0. Master File”), which runs: (1) the “1. Data_Preparation” do-file to generate all variables needed for the analysis; and (2) the “2. Tables and Figures” do-file to produce the final tables and figures. The replicator should expect the code to run in under 30 minutes.

Data and Code Availability Statement:

This paper uses data from the baseline survey of an impact evaluation of vocational training program for women. This dataset is included in this reproducibility package; it is not available in the World Bank Microdata Catalogue.

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 and that the data are anonymized.

Data Sources:

One raw dataset is included in this replication package, consisting of a baseline survey conducted between July 2022 and December 2022 prior to the rollout of the vocational training program. The survey targeted 4,200 eligible women and collected information on demographic characteristics, employment status, and social networks. More details of the data can be found below

Table 1. Data Sources Included in the Replication Package

File name	Location	Description
PHI_IE_Cleaned_PII_community_clean.dta	Working Paper Replication/Data/input_data	Baseline survey dataset used for the analysis.

Software Requirements:

This replication package requires Stata (code was last run with version 18.5). The Stata do-files require the following user-written packages to be installed (the master do-file installs them automatically if needed): mat2txt; estout; lassocoef, esttab, eststo, estpost, and estadd.

Description of code:

This replication package consists of three Stata do-files. All results in the paper can be generated by running a single master script.

The files and their purposes are as follows:

- 0. Master File.do: Sets directory paths, installs required user-written packages (if needed), and executes all steps of the replication. This is the only do-file that needs to be run by the replicator.
- 1. Data_Preparation.do: Constructs social network measures, control variables (as reported in Table 1), and dependent variables used in the analysis. This do-file is called by the master script and does not need to be run separately.
- 2. Tables and Figures.do: Generates figures and tables for the main text and the Annexes using the baseline dataset. This do-file is called by the master script and does not need to be run separately.

The replicator should run 0. Master File.do to reproduce the tables and figures in the paper.

Instructions for replicators:

- Download the dataset in the replication package and place in the '\Data\input_data'
- Edit the file path in line 28 in the '*0. Master File.do*' file to reflect the folder that will be used for replication.
- Run only '*0. Master File.do*' to generate all tables and results.

NOTE: The figures in the paper are constructed in Excel. The code exports the tables in the format required as inputs for figure construction.

Table 2 presents detailed information on the correspondence between tables and figures and their associated replication scripts and outputs.

Table 2-Mapping of Tables and Figures to Replication Code and Outputs

Exhibit name	Output filename(s) (in Tables/ or \$tables)	Script (file + step)	Note (folder, format, etc.)
Table 1 – Summary statistics	Table 1- Summary statistics.xlsx	1. Data_Preparation.do – section computing baseline controls and outcomes	Summary statistics for sample and key controls used in regressions.
Figure 1: Number of people the respondent chat with or talk to, or text, either face-to-face, by phone, internet or any other communication device, on a typical weekday	Figure 1- Number of people.xlsx	2. Tables and Figures.do	Exports the table to Excel, which is then used to generate Figure 1
Figure 2: Proportion of respondents who know someone in key occupations, by type of relationship and number of occupations known	Figure_2.xlsx	2. Tables and Figures.do	Exports the table to Excel, which is then used to generate Figure 2

Figure 3- Networks Occupations by Relationship Type	Figure_3.xlsx	2. Tables and Figures.do	Exports the table to Excel, which is then used to generate Figure 3
Figure 4- Network intensity by Occupation and Women's Economic Empowermen t	Table_4.xlsx	2. Tables and Figures.do	Exports the table to Excel, which is then used to generate Figure 4
Table 2 – Network Size and Women's Economic Empowerment	Table_2_qvals.rtf	2. Tables and Figures.do – “Table 2: Network Size...” block (logit/OLS + BKY q- values)	Main regressions with network size and outcomes (work, paid work, hours, earnings). Exported via esttab to RTF with sharpened q-values.
Table 3 – Network Intensity and Women's Economic Empowerment	Table_3_qvals.rtf	2. Tables and Figures.do – “Table 3: Network Intensity...” block	Main regressions with network intensity and outcomes (work, paid work, hours, earnings). Exported via esttab to RTF with sharpened q-values.
Table 4 – Network Diversity and Women's Economic Empowerment	Table_4_qvals.rtf	2. Tables and Figures.do – “Table 4: Network Diversity...” block	Uses network diversity score (0–14) and exports main coefficients plus BKY q-values.

Table 5 – Network Diversity, Intensity, and Women’s Economic Empowerment	Table_5_qvals.rtf	2. Tables and Figures.do – “Table 5: Network Diversity, Intensity...” block	Regressions with interaction of intensity and diversity.
Table 6 – Exposure to Male-Dominated Occupations and Women’s Economic Empowerment	Table_6_qvals.rtf	2. Tables and Figures.do – “Table 6: Exposure to Male-Dominated Occupations...” block	Uses indicators for knowing family/friend/acquaintance in male-dominated occupations.
Table 7 – Exposure to Female-Dominated Occupations and Women’s Economic Empowerment	Table_7_qvals.rtf	2. Tables and Figures.do – “Table 7: Exposure to Female-Dominated Occupations...” block	Uses indicators for knowing family/friend/acquaintance in female-dominated occupations.
Table 8 – Average Occupational Prestige of Network Resource (SIOPS)	Table_8_qvals.rtf	2. Tables and Figures.do – “Table 8: Average Occupational Prestige...” block	Regressions using average occupational prestige (family, friends, acquaintances) based on SIOPS.
<i>Table A1- Summary statistics: Network indicators</i>	Table A1- Summary statistics_network_indicators.xlsx	2. Tables and Figures.do	Summary statistics for network indicators analyzed throughout the paper.

<i>Table A2- Regression results Network intensity by Occupation</i>	<i>Table A2-Regression results Network intensity by Occupation.xlsx</i>	2. Tables and Figures.do	Each coefficient is estimated in a separate regression.
Annex A4: F- tests results for Equality of Coefficients: Familial Connections vs Diverse Occupational Networks and Occupational Prestige Networks)		2. Tables and Figures.do	
Annex A5: Female and male dominated occupations indicators	Annex A5: Female and male dominated occupations indicators	2. Tables and Figures.do	Summary statistics for female and male dominated indicators.
Annex A7 – Marginal effects for logit models	Table_[i]_me_logit.rtf, i takes the value 2,3, ...8	. Tables and Figures.do – post- estimation marginal-effects section	Contains marginal effects for extensive margin outcomes corresponding to Tables 2–8.
Annex A8 – LASSO and post-LASSO robustness	Table_LASSO_work.rtf ; Table_LASSO_paidwork.rtf ; Table_LASSO_hours.rtf ; Table_LASSO_earnings.rtf	2. Tables and Figures.do – LASSO / post-LASSO block	Robustness tables showing variable selection and post- LASSO estimates for each outcome.