

Read me

Replication files for “Developing a Measure of Women’s Agency in Tunisia: Machine Learning and Generative Artificial Intelligence Approaches” in *World Bank Policy Research Working Paper Series*

August 12, 2026

Questions?

Contact Caroline Krafft, kraff004@umn.edu

Folder structure

In the “Replication TN WAM” folder you downloaded there are various sub-folders: “Expected outputs,” “Do files,” “Data,” “Logs,” and “Graphs and Tables.” Make sure to retain the file structure and subfolders as downloaded.

Instructions for Replicators

Files mostly run out of “TN WAM Root v12 REP” in “Do files.” Note that the root do file includes commented out sections for non-public data preparation (already implemented for public release data, which is the starting point of this analysis).

Note also that because of the use of R, replication is not single shot from just running the root do file; you need to run three distinct pieces, in this order:

1. Run the Stata pieces through line 103
2. Run the R program listed at line 107
3. Return to Stata and run lines 112-end.

You will need to change the directory at the start of the root do file and the setwd in the R program (TN WAM Random Forest v13 CK) to the path for where you saved the folder “Replication TN WAM.”

We attempted to call/install all programs and require the versions we used at the start of do file/R script.

Runs using Stata version 16.1 and R version 4.5.2

Note also that the AI analyses were run using Claude Haiku 4.5, NotebookLM (Google Gemini 3), and Google Gemini 3, as described in the paper. AI is inherently not reproducible; we have video recordings of our prompts (included in appendices) and the AI output (much of which is included in appendices), but the videos cannot be disclosed publicly due to their inclusion of the transcripts. The AI results (e.g., selected questions or GenAI scores) are incorporated into the files we provide.

Data Availability Statement

The **2023 Tunisia MICS** household roster data used for validation are publicly available to registered users from: <https://mics.unicef.org/surveys>

Data citation for Tunisia MICS 2023:

Institut National Statistique (INS) and UNICEF. 2023. Tunisia Multiple Indicator Cluster Survey 2023: Dataset. Tunis, Institut National Statistique (INS) and UNICEF.

Download the files and copy “hl.sav” (the household data) into the “Data/Original data” folder. The analyses uses files downloaded December 19, 2024

You can also choose to comment out line 100 where “TN WAM MICS validation v1 CK.do” is called and skip the MICS (Table 7) replication, which is an appendix table.

The **quantitative survey and manually coded qualitative scores** are publicly available from the World Bank Microdata Library:

<https://microdatalib.worldbank.org/index.php/datadeposit/summary/3214>

Data citation: Krafft, C., Baghdadi, L., Gatti, R., Islam, A.M., Sieverding, M. Forthcoming. Developing a Measure of Women’s Agency in Tunisia: Machine Learning and Generative Artificial Intelligence Approaches. World Bank Policy Research Working Paper.

Download : “TN WAM PUBLIC matched.dta” – This is the main, anonymized data file, which serves as the starting point for the public replication. Please this file in “Data/Analysis data” You will create additional data files through the do files. For instance, the do files then create “TN WAM PUBLIC analysis.dta,” which is the main data set used for analyses.

There are a few additional files that are downloaded with the replication file structure that you will import or merge:

Data/Original data: “LowAI_Claude_allscores_17Apr26.xlsx” – “AI scores” data. Generated using Claude Haiku 4.5 on April 17, 2026 using an institutional license through the web interface.

Data/Original data: “Low-mid AI_Claude_all scores_20Apr26” – “AI domains” data. Generated using Claude Haiku 4.5 on April 20, 2026 using an institutional license through the web interface.

Data/Original data: “TN WAM variable text.xlsx” – text for the questions asked in the survey

Data/Original data: “AI text only.xlsx” – the questions created by the two AI methods that are text generation. AI qual.- AI quest. portion of data from NotebookLM (Google Gemini 3) using a web interface and institutional license on March 10, 2026. AI gen. quest portion from Google Gemini 3, using a web interface, March 10, 2026.

Data/Original data: “JBC.data” a dataset with the variables selected by the Jayachandran et al. (2023) paper. Specifically, results are from the paper available at:

<https://doi.org/10.1016/j.worlddev.2022.106076>

Output (List of Exhibits)

Output will run out into “Graphs and Tables.” The folder “Expected outputs” includes copies of all the raw outputs that you should obtain.

Tables

Table 1: Comparison of GenAI model designs based on inputs. This is a text-only construction by the authors describing the features of the different AI.

Table 2: Questions selected by ML and GenAI by method

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/ranked_all_top5_method.xlsx

Note columns are slightly reordered in paper.

Table 3: Questions selected by Tunisia vs. India context, overlapping question set

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/ranked_top5_country.xlsx

Table 4: Questions selected by rural versus urban context

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/ranked_top5_location.xlsx

Table 5: Questions selected by option set – DHS and LMPS

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/ranked_top5_LMPS_DHS.xlsx

Table 6: Questions selected by number of observations, LASSO stability selection models

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/ranked_top5_lasso_Ns.xlsx

Table 7: Comparisons between WAM Survey and MICS: Means, standard errors, differences, and pairwise t-tests

Runs out of “TN WAM MICS validation v1 CK.do”

File: Graphs and Tables/WAM MICS valid.xlsx

Table 8: Correlation between different qualitative scores and domains, by methods

Runs out of “TN WAM descriptives v5 CK.do”

File: Graphs and Tables/AI_corr.csv

Table 9: Standardized qualitative scores and index of all quantitative scores, by method and demographics

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/different qual by demo.xls

Figure 1: Manual scoring method histogram of domain qualitative score distributions (percentages)

Runs out of “TN WAM descriptives v5 CK.do”

Complication of files: Graphs and Tables/dist_*.png where * is variously c_qual_education c_qual_health c_qual_ave_mob c_qual_ave_fert c_qual_ave_fin c_qual_emp

Figure 2: AI scores histogram of domain qualitative score distributions (percentages)

Runs out of “TN WAM descriptives v5 CK.do”

Complication of files: Graphs and Tables/dist_*.png where * is variously c_low_education c_low_healthcare c_low_employment c_low_ave_mob c_low_ave_fert c_low_ave_fin

Figure 3: AI domains histogram of domain qualitative score distributions (percentages)

Runs out of “TN WAM descriptives v5 CK.do”

Complication of files: Graphs and Tables/dist_*.png where * is variously c_mid_HHDomainAvg c_mid_EconDomainAvg c_mid_Physical_Mobility c_mid_Social_Cultural

Figure 4: Manual scoring method histogram of average overall qualitative score (percentage)

Runs out of “TN WAM descriptives v5 CK.do”

File: Graphs and Tables/dist_c_qual_ave_dim.png

Figure 5: AI scoring histogram of average overall qualitative score (percentage)

Runs out of “TN WAM descriptives v5 CK.do”

File: Graphs and Tables/dist_c_low_ave_dim.png

Figure 6: AI domains histogram of average overall qualitative score (percentage)

Runs out of “TN WAM descriptives v5 CK.do”

File: Graphs and Tables/dist_c_mid_ave_dim.png

Figure 7: Correlation coefficient for manual qualitative score with standardized index of five questions selected by model

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/correlation plot.png (corresponding numerical values in Graphs and Tables/correlation matrix.csv)

Figure 8: Correlation coefficient for manual qualitative score with five random questions over 1,000 random iterations

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/correlation random.png

Figure 9: Correlation coefficient for qualitative scores, by domain and method, with standardized index of all 89 quantitative questions

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/quant_all_corr.png (corresponding numerical values in Graphs and Tables/quant_correlation matrix.csv)

Figure 10: Correlation coefficient for manual qualitative score by number of questions selected, LASSO stability selection models

Runs out of “TN WAM combined outputs v9 CK.do”

File: Graphs and Tables/no_quest SA.png