Read Me

Replication Package

10-02-2025

Paper: Does perceived labor market competition increase prejudice between refugees and their local hosts? Evidence from Uganda and Ethiopia.

Authors: Julie Bousquet; Anna Gasten, Mark Marvin Kadigo, Jean-François Maystadt, Colette Salemi

Contact: Colette Salemi at csalemi@uvic.ca

# RoadMap

## Folders

The replication package includes four folders

* **01. Datasets**, which includes
  + the “base data” in *01. Base*, which includes the raw data used for our analysis.
  + The “temporary data” in *02. Temp*, which includes any temporary datasets created through the analysis.
  + The “final” folder in *03. Final*, which includes the final dataset, created from the cleaning do file 01\_EXP\_MergeClean.do.
  + The “LSMS data” in *04. LSMS*, which includes the used LSMS data for some robustness checks as part of our analysis. The LSMS data is also publicly available through the World Bank Microdata Library: The [Ethiopian Socio-Economic Panel Survey 2021-2022](https://microdata.worldbank.org/index.php/catalog/6161); and the [Ugandan National Panel Survey 2019-2020](https://microdata.worldbank.org/index.php/catalog/3902).
* **02. Output**, which includes
  + the replicated paper “*Main*” in latex format, and
  + the “*Figures*” and “*Tables*” folders, which include the tables and figures.
* **03. Codefiles** with more detailed explanations below
  + Master do file *00\_EXP\_MasterDoFile*.do
  + Cleaning do file *01\_EXP\_MergeClean*.do
  + Descriptives do file *02\_EXP\_Descriptives*.do
  + Analysis do file *03\_EXP\_Analysis*.do, which uses the correction for multiple hypothesis testing *2a.fdr\_qvalues*.do
  + Heterogenous do file *04\_EXP\_Heterogeneity*.do
  + Robustness do file *05\_EXP\_Robustness*.do
  + Power Calculation R file *06\_EXP\_Power calculations*.rmd
* **04. Documentations** with
  + 01. Questionnaire, including the *Household questionnaire* and the *RSI questionnaire* (Randomly Selected Individual). The experimental module can be found in the *RSI Questionnaire* document in Section 9.
  + 02. PAP (pre analysis plan)

## Instructions

To replicate the paper's figures and tables, the **master do-file** must be opened first. Users should enter their username and uncomment the indicated line (/\*) in the do-file to run all scripts simultaneously.

As long as the folder structure remains unchanged, the replication should proceed smoothly. However, some packages may need to be installed or enabled if required.

The **power calculation R file** can be executed separately without modifying any file paths. It is self-contained and generates the relevant table in a single run.

### Key variables of interest

* Unique ID: ciid
* Localities: region\_short
* Weight: w
* Treatment group: treat\_var
* Cluster: cluster\_psu
* Refugee status: refugee
* InGroup: exp\_ingroup
* Same Occupation: exp\_same\_occup
* Outcome: prej\_aind\_byR

## Structure

### Datasets

The key raw datasets used for the analysis are:

* Ethiopia
  + Household: ETH-HH\_Main\_User\_Weights.dta ;
  + Roster: ETH-Roster\_User\_Weights.dta ;
  + Randomly selected individual: ETH-RSI\_Main\_User\_Weights.dta ;
* Uganda
  + Household UGA-HH\_Main\_User\_Adaptive W.dta and UGA-HH\_Main\_User\_Weights.dta ;
  + Roster UGA-Roster\_User\_Adaptive W.dta and UGA-Roster\_User\_Weights.dta ;
  + Randomly selected individual: UGA-RSI\_Main\_User\_Adaptive W.dta and UGA-RSI\_Main\_User\_Weights.dta

### Code Files

#### \*Preparation/Cleaning Do File\*

01\_EXP\_MergeClean.do

Creates the final dataset *04\_UGA\_ETH\_Prepared*.dta

#### \*Descriptives\*

02\_EXP\_Descriptives.do

This do file generates

* Balance tables: Table A7: TA7\_Balance\_Tables\_All.tex; TA7\_Balance\_Tables\_Hosts.tex; TA7\_Balance\_Tables\_Refugees.tex are manually aggregated into table TA7\_Balance\_Tables.tex.
* Summary Statistics
  + Table A8: TA8\_SummStat\_Uganda.tex
  + Table A9: TA9\_SummStat\_Ethiopia.tex

#### \*Analysis\*

03\_EXP\_Analysis.do

This do file generates

* Main analysis: Table A11: TA11\_Analysis\_PrejudiceIndex.tex
* Main analysis: Figure 4: FIG4\_CFP\_Analysis.pdf
* By localities: Table A16: TA16\_Analysis\_byLocality.tex
* By localities: Figure 5: FIG5\_CFP\_Analysis\_byLocalities.pdf

Note: the 2a.fdr\_qvalues.do do file is used to correct for Multiple Hypothesis Testing.

#### \*Heterogeneity\*

04\_EXP\_Heterogeneity.do

This do file generates

* By gender and education: Table A18: TA18\_Analysis\_Gender\_Education.tex
* By Over-work: Table A19: TA19\_Analysis\_OverWork.tex
* By Occupation Pressure: Table A20: TA20\_Analysis\_OccupationPressure.tex
* By network/contact: Table A21: TA21\_Analysis\_Contact.tex
* By language shared with in group: Table A22: TA22\_Analysis\_LanguageShare.tex
* By language fractionalization: Table A23: TA23\_Analysis\_LanguageFraction.tex
* Table 1: TAB1\_Analysis\_HeterogMutli.tex ; Aggregates for hosts of tables
  + By Over-work: Table A19
  + By Occupation Pressure: Table A20
  + By network/contact: Table A21
  + By language shared with in group: Table A22

#### \*Robustness\*

05\_EXP\_Robustness.do

This do file generates

* Comparison summary statistics with LSMS: Table A4: TA4\_Summary\_Statistics\_LSMS.tex
* By components of the index: Table A12: TA12\_Analysis\_byIndex.tex
* Labor Market competition: Table A10: TA10\_Analysis\_LMCompetitionIndex.tex
* Dummy Variable: Table A13: TA13\_Analysis\_DummyIndex.tex
* Without clustered SEs: Table A14: TA14\_Analysis\_ClusteredSEs.tex
* By a set of employment variables: Table A15: TA15\_Analysis\_EmploymentDef.tex
* Regression Coefficient Comparisons Across Locations (diff-in-mean): TA17\_Analysis\_PValues.tex.

#### \*Power Calculation\*

This Rfile 06\_EXP\_Power calculations.rmd generates Table A24: TA24\_Retrospective\_PC.tex