READ-ME: Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone

Overview

This is a replication package for Social Signaling and Childhood Immunization: A Field Experiment in Sierra Leone by Anne Karing.

The package contains the code and data files necessary to replicate the analysis presented in Karing (2024). Three main do files run all of the code to generate the data for the 10 figures and 31 tables in the paper.

Software: Replication requires Stata Version 18 or later and Matlab Version 22 or later.

Runtime: The replicator should expect the code to run for between 45 minutes to 2 hours.

Folder structure:

- **Documentation**
  - README.pdf
  - **Surveys**
- **Data**
  - **Intermediate**
  - **Final**
- **Dofiles**
  - MasterDofile_Filepaths.do
  - MasterDofile_Tables.do
  - MasterDofile_Figures.do
  - iebaltab.ado
  - **Tables**
  - **Figures**
  - **Appendix**
    - **ATables**
    - **AFigures**
- **matlab**
- **Output**
  - **Tables**
  - **Figures**
  - **Appendix**
    - **Tables**
    - **Figures**

Data Availability and Provenance

The data were generated through surveys conducted between 2016 and 2021, as part of a field experiment implemented by Anne Karing from 2016 to 2018. For more information on the different data sources and how they were obtained, see Section III.E of the paper.

Raw data is used to construct the data sets for analysis. This process involves merging different surveys on personally identifiable information (PII) variables. The data provided in this package contains the data set prepared for analysis after all PIIs are dropped to preserve respondents' confidentiality.

Statement about Rights

- ✓ I certify that the author of the manuscript has legitimate access to and permission to use the data used in this manuscript.
- ✓ I certify that the author of the manuscript has documented permission to publish the data contained within this replication package.

Data sets:

The data sets stored under Data/Intermediate are included as base data sets, meaning they are not generated by running the replication code. Below is a description of each of them including how they are constructed.

- **Balance_ClinicCommunityCharacteristics.dta**

  This data set is constructed at the clinic- or community-variable level using four sources: i) surveys with nurses in charge of study clinics at baseline, ii) surveys with community members and stakeholders implemented at the start of the experiment during information meetings; iii) monitoring forms used to collect data on the implementation during the experiment, and iv) surveys with mothers at the exit of study clinics during the experiment. It is used to check for balance of clinic and community characteristics, including clinic services provided and perceptions of community members about immunization, and monitoring intensity across control and treatment groups (Table B2).

- **Balance_StudyEndlineSample.dta**

  This is an individual-level data set created using the de-identified endline survey data. It contains all variables related to socio-demographic characteristics of the mother and is used to check for balance across control and treatment groups (Table B1).

- **Balance_ListingSample.dta**

  This is an individual-level data set created using the de-identified listing and follow-up survey data. It contains variables related to mothers' residence status in the community and characteristics of the children included in the main analysis. It is used to check balance and attrition across control and

treatment groups for the full sample (Table B3) as well as the 24-month sample (Table B4).

- **Endline_Survey_Data.dta**

  This is an individual-level data set created using the de-identified endline survey data. It contains variables related to mothers' knowledge, beliefs and preferences around vaccination (Figures IV and A6 and Tables VI, B8, B9, B10, B11 and B18).

- **Endline_Distance_Analysis.dta**

  This is an individual-level data set created using three different sources: i) the endline survey data containing detailed socio-demographics, ii) the listing and follow-up survey data containing vaccination behavior, and iii) the community-level data on distance between communities and clinics. This data set is used for the analysis of the relationship between distance, socio-demographic characteristics, and vaccination behavior and treatment effects (Tables B19 and B20).

- **FirstOrderBeliefs_Data.dta**

  This data set is at the individual-mother level and is created using the de-identified endline survey data and data on vaccination status of children from the latest administrative clinic records. It is used for the analysis of respondents' knowledge of other mothers' vaccination behavior and their child's bracelet (Tables I, II and III).

- **FirstOrderBeliefs_Constructed.dta**

  This data set is constructed using the same sources as FirstOrderBeliefs_Data.dta. It is used to display mothers' beliefs about other children's vaccine status based on their bracelet color to create Panel A of Figure III.

- **FirstOrderBeliefs_Truth_Constructed.dta**

  This data set is constructed using the same sources as FirstOrderBeliefs_Data.dta. It is used to compare mothers' beliefs about other children's vaccination status with the truth to create Panel B of Figure III.

- **Listing_Implementation.dta**

  This is an individual-level data set created using the de-identified listing survey data. It contains variables related to bracelet implementation and is used for the analysis on correct handout of bracelets by clinics (Figure A5 and Table B5).

- **Listing_Data.dta**

  This is an individual-level data set created using the de-identified listing survey data. It contains variables related to bracelet ownership, and wearing and retention behavior (Figure II and Table B17).

- **ListingFollowUp_analysis.dta**

  This is an individual-level data set created using the de-identified listing and follow-up survey data. It contains all variables related to vaccination behavior and is used for all analyses on vaccination outcomes (Figures V and A9, Tables IV, V, B12-16, B22-23).

- **ListingFollowUp_TreatmentEffectsOverTime.dta**

  This data set is constructed using ListingFollowUp_analysis.dta as a base and contains variables related to vaccination behavior constructed for the estimation of cohort-specific treatment effects (Figure A7 and A8).

- **ListingFollowUp_analysis_attrition.dta**

  This is an individual-level data set created using the de-identified listing and follow-up survey data. It contains variables related to vaccination behavior as well as children's attrition status between the listing and follow-up surveys. It is used to analyze the effects of the treatment by whether a child was surveyed during the follow-up (Table C1)

- **MasterData_PilotClinics.dta**

  This is a clinic-level data set containing indicator for whether the clinic was part of a pilot study. It is used across a number of analyses.

- **MasterData_CommunityDistances.dta**

  This is a community-level data set containing information on each community's distance from the clinic. It is used across a number of analyses.

- **MasterData_LaunchDates.dta**

  This is a clinic-level data set containing information on the date at which the program started in each clinic. It is used across a number of analyses.

- **SecondOrderBeliefs_Data.dta**

  This data set is at the individual-mother level and is created using the de-identified endline survey data. It contains all observation relevant to the analysis of respondents' belief about other mothers' knowledge of their vaccination behavior (second-order beliefs).

The following data sets stored under Data/Final are included as generated data sets, which means they are recreated each time the replication code is executed:

- Endline_Externalities_constructed.dta
- KnowledgeTable_FirstOrderBeliefs_analysis.dta
- listing_Arm3_AgeCat_GreenBracelets.dta
- listing_Arm4_AgeCat_GreenBracelets.dta
- Listing_BraceletHandout_constructed.dta
- Listing_HandOutGreenBracelets_S4_S5_prepared.dta
- Listing_Data_BraceletHandout.dta

- ListingFollowUp_byAgeOneYear_analysis.dta
- ListingFollowUp_ExtensiveMargin_analysis.dta
- ListingFollowUp_timely_analysis.dta
- SICI_FollowUp_BirthCohort_Variable.dta

Replication Instructions

Software requirements

- Stata (code was last run with Version 18) packages:
  - ietoolkit
  - fre
  - texsave
  - dataout
  - matnames
  - distinct
  - randcmd
  - colrspace
  - binscatter
  - blindschemes
  - grc1leg
  - grstyle
  - estout
  - palettes
- Matlab (code was last run with Version 24.1)
  - Optimization Toolbox
  - Statistics and Machine Learning Toolbox

Memory and Runtime Requirements

Approximate time needed to reproduce the analyses on a standard 2020 desktop machine: 45 min - 2 hours.

Details: The code was last run on a M1-chip based laptop with MacOS version 11.7.10.

Instructions to Replicators

- Edit Dofiles/MasterDofile_Filepaths.do to adjust the default path.
- Run Dofiles/MasterDofile_Filepaths.do to download the necessary packages and set analysis parameters such as seed and number of replications.
- Run Dofiles/MasterDofile_Figures.do to produce all manuscript and appendix figures.
- Run Dofiles/MasterDofile_Tables.do to produce all manuscript and appendix tables.
- All output generated will be located in Output.

- All programs can be run in parallel, except for the do file generating Appendix Table B5 which needs to be run after that generating Figure II.

Structural estimation: In order to reproduce Table E1, replicators need to use the files located in the folder titled Matlab and take the following steps:

- Open Matlab and import the following files and datasets:
    1. estimate_dynamic_UI.m
    2. estimate_dynamic.m
    3. likelihood_objective.m
    4. likelihoods.m
    5. dynamic_data_2022_UI.csv
    6. dynamic_data_2022.csv
- Run estimate_dynamic.m and estimate_dynamic_UI.m.
- Drop the output structuralmodel.csv and structuralmodel_UI.csv in the Data/Intermediate folder.
- Run the do file Appendix/ATables/Table_StructuralEstimation.do.

Controlled Randomness

Random seed is set at line 56 of program MasterDofile_Filepaths.do.

All files are meant to be replicated.

List of tables and programs

The provided code reproduces:

Figures:

- Figure II: Correct Hand Out of Bracelets by Treatment Groups. File path and name: Figures/Figure_CorrectHandOutBracelets.do
- Figure III: Parents' Inferences about other Children's Vaccinations From Bracelet Colors. File path and name: Figures/Figure_Tables_EffectsOfSignalsOn_Beliefs_Truth_Combined.do
- Figure IV: Reputational Costs and Benefits. File path and name: Figures/Figure_ReputationalCostsBenefits.do
- Figure V: The Effect of Distance to the Clinic on Timely Vaccination. File path and name: Figures/Figure_DistanceTakeUp_Timeliness_ControlvsS5-_binscat.do

Tables:

- Table I: The Visibility of Bracelets by Treatment Group. File path and name: Tables/Table_VisibilityOfBraceletsByTreatmentGroup.do

- Table II: The Effects of Signals on Second-Order Beliefs and Knowledge of Other Children's Vaccinations. File path and name: Tables/Table_FirstSecondOrderBeliefs_VaccineDecisions.do
- Table III: The Effects of Signals on Knowledge of Other Children's Vaccinations by Bracelet Color. File path and name: Tables/Table_FirstSecondOrderBeliefs_VaccineGuessing_SignalingPeriod.do
- Table IV: The Effects of Signals on Timely and Complete Vaccination, Separated by Treatment. File path and name: Tables/Table_Panel_Timely_byAgeOneYear.do
- Table V: The Effects of Signals on Complete Vaccination by 12, 18, and 24 months. File path and name: Tables/Table_EffectsOfSignals_ExtensiveMargin_Vaccine4and5.do
- Table VI: The Effect of Signals on Preferences for Different Vaccinations. File path and name: Tables/Table_PreferencesVaccinations.do

Appendix Figures:

- Figure A5: Hand Out of Green Bracelets in Signals at 4 and 5 according to Timely Vaccination. File path and name: Appendix/AFigures/Figure_HandOutOfGreenBracelets_S4_S5_TimelyVaccination.do
- Figure A6: Reputational Costs and Benefits in terms of Actions. File path and name: Appendix/AFigures/Figure_ReputationalCostsBenefits_Actions.do
- Figure A7: Treatment Effects by Birth Cohorts for Timely Vaccination - Vaccine 4. File path and name: Appendix/AFigures/Figure_TreatmentEffectsOverTime_Vaccine4_byCohorts.do
- Figure A8: Treatment Effects by Birth Cohorts for Timely Vaccination - Vaccine 5. File path and name: Appendix/AFigures/Figure_TreatmentEffectsOverTime_Vaccine5_byCohorts.do
- Figure A9: The Effect of Distance to the Clinic on Timely Vaccination. File path and name: Appendix/AFigures/Figure_DistanceTakeUp_Timeliness_UIvsS5_binscat.do

Appendix Tables:

- Table B1: Characteristics of Endline Survey Sample. File path and name: Appendix/ATables/BalTable_EndlineSample.do
- Table B2: Description of Clinic and Community Characteristics. File path and name: Appendix/ATables/BalTable_ClinicCommunityCharacteristics-.do
- Table B3: Characteristics of Main Analysis Sample. File path and name: Appendix/ATables/BalTable_ListingSample.do
- Table B4: Characteristics of 24 Months Analysis Sample. File path and name: Appendix/ATables/BalTable_FlwupSample_at24m.do
- Table B5: Verifying the Correct Implementation of Bracelets, Regression Results for "Figure II". File path and name: Appendix/ATables/Table_CorrectImplementation.do

- Table B6: How Mothers Know about the Color of the Bracelet. File path and name: Appendix/ATables/Table_KnowledgeColorBracelet.do
- Table B8: Reference Groups for Social Signaling. File path and name: Appendix/ATables/Table_ReferenceGroups.do
- Table B9: Private and Social Benefits of Vaccination, Knowledge of Externalities. File path and name: Appendix/ATables/Table_Externalities.do
- Table B10: Others' Inferences about Types Conditional on Vaccination Decisions. File path and name: Appendix/ATables/Table_OthersInferencesAboutTypes.do
- Table B11: Inferences about Types Conditional on Vaccination Decisions. File path and name: Appendix/ATables/Table_InferencesAboutTypes.do
- Table B12: The Combined Effects of Signals at 4 and 5 on Timely and Complete Vaccination, Separated by Treatment. File path and name: Appendix/ATables/Table_Panel_CombinedEffectsS4S5.do
- Table B13: Effects of Signals on Timely Vaccination Using Different Cutoffs. File path and name: Appendix/ATables/Table_TimelyVaccination_Robustness.do
- Table B14: Effects of Signals on Timely- and Complete Vaccination by Age One Year (Variable Sample). File path and name: Appendix/ATables/Table_EffectsOfSignals_Timely_AgeOneYear_WithControls_VariableSample.do

The provided code does not generate the following content in the paper:

- Figure I: Experimental Treatment Groups
- Figure A2: Different Bracelets handed out across the Three Signaling Treatments
- Figure A2: Clinic Randomization
- Figure A3: Process of Community Selection
- Figure A4: Final Sample Flow Diagram
- Table B21: Cost-effectiveness of Interventions to Increase Childhood Vaccination