Readme File for Replication Package of
"Religious Divisions and Production Technology: Experimental Evidence from India"
Arkadev Ghosh
17 December 2024

OVERVIEW

This README file describes the contents of the replication folder and instructs users how to replicate the results of the paper and appendix. Separate files to reproduce the tables and figures are stored in the /Code/ folder.

To perform the replications, run "00 main.do" in the code folder. Before doing so, change the global projectfolder "dir" to the location of the Replication folder on your computer.

The replication package includes all code (cleaning and analysis) needed to generate the results in the paper and appendix.

DATA AVAILABILITY AND PROVENANCE

The data used in this manuscript were collected/accessed with full permission and in compliance with agreements made with the partnering organization. Most of the data are confidential and protected under a Non-Disclosure Agreement (NDA) -- the author does not have the rights to distribute this. The data on individual-level productivity, beliefs about productivity, and survey of firm managers are exempt from these restrictions and are included in the provided package.

DATA FILE DESCRIPTIONS

The following data files produce all figures and tables in the paper and appendix.

Folder Data/Original

1. `Time Use Data.csv` - This file includes stopwatch measures of coordination or dependency among workers at the line-section-job level collected by RAs (some sections have more than one job).
2. `baseline_raw.dta` - This file includes raw baseline survey data at individual-worker level.

3. employee_task_history.dta - This file includes employees' task history data, their first job in the factory and their last job before the intervention, and any other job held for more than 6 months.
4. teamlist_baseline.dta - This file includes the line-section-shift level assignments of workers at baseline from the HR manager.
5. Baseline_LS_composition.dta - This file contains statistics on the religious composition at the line-section level at baseline.
6. Randomized_Teams.dta: This file includes post-randomization team assignments.
7. Randomized_Teams_RI.dta - This file includes 2000 alternate assignments for each worker, used for randomization inference.
8. worker_attitudes_endline.dta - This file includes (raw) worker attitudes data from the endline survey at individual-worker level.
9. worker_teammate_interactions.dta – Raw worker-teammate-interaction level (Free Riding, Blame, Relief time etc.) data from the endline survey.
10. attrited.dta - This file includes the attrition data at individual-worker level obtained from the HR.
11. beliefs_raw.dta - This file includes the raw data of individual beliefs from the paper bag making task.
12. individual_productivity_raw.dta - This file includes raw data on individual productivity from the paper bag making task.
13. individualprod_ids_religion.dta - This file includes religion indicators for participants in the paper bag making task from a baseline survey conducted with them.
14. slot_schedules.dta - This file includes the schedule of slots for the 10 waves of data collection for the paper bag making task.
15. timeline_individualdata.csv - This file contains the timeline details for data collection from the paper bag-making task.
16. firm_manager_survey.dta - This file includes raw survey data completed by managers at different factories.

In addition to the datasets above, there are a few additional datasets in the Data/Original Folder that are created from the datasets above.

17. line_section_dependency_mins.dta - This file includes coordination time (dependency) data at line-section level.
18. linesection_sharemuslims.dta - This file includes the share of Muslims in each line-section team after randomization.
19. PropOldTeammates.dta - This file contains data for Hindu workers in mixed teams, showing the proportion of Muslim workers (in their new teams) who were also part of their team prior to randomization.
20. Section_Aggregates.dta – This file includes team characteristics data at line-section-team level post randomization used for the various heterogeneity tests at the line-section-level.

Folder Data/Final

1. Balance_check.dta - This file includes baseline in-person survey data at individual-worker-level merged with randomized teams' data.
2. Baseline_Cleaned.dta – This file includes cleaned baseline in-person survey data at individual-worker level.
3. beliefs_cleaned.dta - This file includes cleaned data on beliefs about ingroup and outgroup productivity from the paper bag making task.
4. individual_produc_clean.dta - This file includes cleaned individual productivity data with religion indicators from the paper bag making task.
5. linelevel_data.dta - This file includes output data at line level obtained from the firm.
6. linesectionlevel_data.dta - This file includes output data at line-section-level obtained from the firm.
7. linesectionlevel_datamerged.dta – This file includes line-section-level data merged with section level characteristics (Section_Aggregates.dta) for heterogeneity analysis.
8. linesectionlevel_starbility.dta - This file includes line-section-level data for robustness tests using the Starbility package in R.
9. worker_attitudes.dta - This file includes worker attitudes data from the endline survey cleaned and merged with baseline data.
10. worker_choices.dta - This file includes worker-teammate-choice level data cleaned and merged with baseline data.

COMPUTATIONAL REQUIREMENTS

The code was last run using Stata Version SE 18 on Windows 11 with packages updated on December 17th, 2024. The code was also run successfully using Stata SE 16 on Mac OS 12.5. It should run on relatively recent versions of Stata in Windows and Mac OS, but we cannot guarantee that it will. All required Stata packages are loaded in the master.do file. No other software is required.

The code was last run on a machine with the following specifications:

- Processor: AMD Ryzen Threadripper PRO 5955WX 16-Cores 4.00 GHz
- Installed RAM: 128 GB (128 GB usable)
- System type: 64-bit operating system, x64-based processor

The code takes approximately 22 minutes to run from start to finish on the above machine.

CODE

The 01 Prep folder includes all cleaning code and 02 Analysis includes all of the Analysis code.

Folder: 01 Prep

1. `00 dependency summary stats.do` – Prepare dependency (mins) data at line-section level from time-use data.
2. `01 prep balance.do` – Prepare the baseline survey data for analysis, and merge cleaned baseline data with randomized teams for balance checks.
3. `02 section_aggregates.do` - Prepare baseline survey data, and merge cleaned baseline data with randomized teams' data and the share of Muslims in each line-section-team after randomization to generate section aggregates data at line-section-team level for heterogeneity analysis at the line-section-level.
4. `03 worker_interactions_clean.do` - Prepare worker-teammate-choice level data (blame, free riding, relief time) merged with cleaned baseline data for analysis.
5. `04 worker_attitudes_clean.do` – Prepare worker attitudes data from the endline survey and merge it with cleaned baseline data for analysis.

6. `05 individual productivity clean.do` - Prepare individual productivity data from the paper bag making task and merge it with the religion indicator for analysis.
7. `06 beliefs_clean.do` - Prepare baseline individual belief data from the paper bag making task and merge it with the religion indicator for analysis.

Folder: 02 Analysis

1. `01 dependency summary stats.do` – Prepare dependency (mins) data at line-section level from time-use data; Table 1, Table B.2, Figure 3.
2. `02 balance_tables.do` - Table A.4, Table A.3, Table A.5, Table A.8, Table B.12, Table A.7, Table A.6, Table B.1, Table B.14, Table A.9, Table A.1, Table A.2, Figure C.9.
3. `03 line_level_regressions.do` - Table 2, Table B.3, Table B.4, Figure 7, Table B.17
4. `04 linesection_level_regressions.do` - Table 3, Table B.7, Table B.8, Table B.5, Table B.13. Table B.10, Table 6, Table B.18, Table B.19, Figure C.3, Table B.6, Figure C.4, Figure C.5, Table B.24, Table B.23, Table B.9.
5. `04a starbility.Rmd` - Figure C.2.
6. `05 worker_teammate_regressions` - Table 4, Table B.11, Table B.15, Table B.20.
7. `06 worker_level_regressions.do` - Table 5, Table B.16, Table B.21.
8. `07 allocation_comparisons.do` - Figure 2
9. `08 belief_analysis.do` - Figure C.7
10. `09 indiv_produc_analysis.do` - Figure C.6
11. `10 firm_manager_survey.do` - Figure C.14, Figure C.15, Figure C.16.
12. `11 output_ratings_correlation.do` – Figure C.10, Figure C.11, Figure C.12
13. `12 survey_quality.do` – Figure C.13