

README for the Reproducibility Package for How Do Organizations Learn? The Diffusion of Scientific Evidence on Generative AI

Overview

This replication package contains the code used to reproduce the results of the impact evaluation on information diffusion within the World Bank. The analysis combines survey and administrative data from a field experiment that studies how scientific evidence on the impacts of generative AI spreads within a large organization. The package reproduces all tables and figures in the paper, including experimental impacts of sharing research findings with senior versus junior staff and tests of belief-based mechanisms related to peer adoption and evidence credibility.

Data Availability Statement

Dataset Name	File Name(s)	Link/Access Instructions	Access Date	Data Access
Wave 1 Seed Data	Wave 1 Seed.dta	Survey data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 2 Seed Data	Wave 2 Seed.dta	Survey data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 1 Contact Data	Wave 1 Contact.dta	Survey data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 2 Contact Data	Wave 2 Contact.dta	Survey data. The data will be available in the MicroData library in the future	2025	Restricted
Staff position indicators	Position data.dta	Administrative data. The data is restricted. Please reach out to	2025	Restricted

		mshaukat@worldbank.org.		
Wave 1 Treatment	Wave 1 Treatment.dta	Randomization data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 2 Treatment	Wave 2 Treatment.dta	Randomization data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 1 Article Clicks	wave1_article clicks.csv	Bitly data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 1 Infographic Clicks	wave1_infographic clicks.csv	Bitly data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 2 Article Clicks	wave2_article clicks.csv	Bitly data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 2 Infographic Clicks	wave2_infographic clicks.csv	Bitly data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 1 Seed Batches	Wave 1 Seed_batch.dta	Randomization data. The data will be available in the MicroData library in the future	2025	Restricted
Wave 2 Seed Batches	Wave 2 Seed_batch.dta	Randomization data. The data will be available in the	2025	Restricted

		MicroData library in the future		
Wave 1 URL Links	Wave 1 Links_withNewID.dta	URL data. The data is restricted. Please reach out to mshaukat@worldbank.org.	2025	Restricted
Wave 2 URL Links	Wave 2 Links_withNewID.dta	URL data. The data is restricted. Please reach out to mshaukat@worldbank.org.	2025	Restricted
Base Dataset	Base Dataset.dta	Administrative data. The data is restricted. Please reach out to mshaukat@worldbank.org.	2025	Restricted
Expert Predictions SSPP	Evidence Transmission_SSPP.xls	Expert prediction data. The data will be available in the MicroData library in the future	2025	Restricted
Expert Predictions World Bank	Evidence Transmission_WP.xls	Expert prediction data. The data will be available in the MicroData library in the future	2025	Restricted

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.

I certify that the author(s) of the manuscript have documented permission to redistribute/publish the data contained within this replication package. Appropriate permissions are documented in the LICENSE.txt file.

Instructions for Replicators

1. Update the global macro \$mainFolder to the root directory where your data and output folders reside.

2. Run the Master.do file. This file (i) prepares data; and (ii) outputs all tables and figures in the main paper and appendix.

Code Description

Data Preparation

- **Prepare Bitly data:** Load, clean, and construct outcome measures capturing engagement with study materials
- **Prepare SEED data:** Clean and process survey data on beliefs, credibility perceptions, and knowledge of research findings. Merge survey and administrative datasets to create the analysis sample used in the experiment.
- **Prepare CONTACT data:** Clean and process survey data on recall of research findings. Merge survey and administrative datasets to create the analysis sample used in the experiment.

Main Analysis

- **Main Tables 1–3:** Estimate experimental impacts of sharing research findings on information transmission, engagement, and recall outcomes using survey data.

Appendix A

Figures

- **Figure A2:** Generate figures summarizing expert predictions of treatment effects.
- **Figures A3–A7:** Produce descriptive figures on engagement, diffusion patterns, and staff characteristics.
- **Figure A8:** Estimate and plot heterogeneity in treatment effects across subgroups.

Tables

- **Table A3:** Estimate long-term effects of the experimental interventions.
- **Tables A4–A5:** Re-estimate main results excluding managers.
- **Tables A6–A7:** Re-estimate main results restricting to staff in policy-relevant positions.
- **Table A8:** Analyze experimental impacts on ChatGPT adoption outcomes.
- **Tables A9–A10:** Present additional heterogeneity analyses.
(Note: Appendix Tables A1–A2 rely on confidential data and cannot be fully replicated with the public files.)

Appendix B

- **Figures B1–B4:** Create descriptive figures on ChatGPT usage and engagement with evidence based on survey data.

Computational Requirements and Runtime

Runtime is 2-3 minutes.

Software Requirements

Stata version 15 or higher (for running the .do files)