README File
Replication Package for "Selection, Absolute Advantage, and the Agricultural Productivity Gap"

Francisco Alvarez-Cuadrado[§] Francesco Amodio[¶] Markus Poschke[‖]

May 2025

1 Overview

This replication package contains the code and data to replicate all Tables and Figures in Alvarez-Cuadrado, Amodio and Poschke (2025) "Selection, Absolute Advantage, and the Agricultural Productivity Gap," *Journal of Political Economy Macroeconomics*. The package is deposited in the JPE Macro Dataverse.

2 Data Availability Statement

"*We certify that the authors of the manuscript have legitimate access to and permission to use the data used in this manuscript, and have permission to redistribute/publish the data contained within this replication package.*"

3 Data

The paper uses data from the World Bank's Living Standards Measurement Study - Integrated Surveys on Agriculture (LSMS-ISA) project. The data are available online at `https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA`. We downloaded them on October 9, 2018 and make the raw data available in this replication package in the `./Replication/LSMS-ISA` folder.

Simulated model data plotted in Figures 1, 2, 4 and 6 have been generated with the Matlab code file `generate_data.m` in the `model` folder, and are saved in the files `F3_data.mat`, `F1a_Ag_misaligned_data.mat` and `F1a_Ag_aligned_data.mat`.

4 Computational and Run-time Requirements

- **Stata**

[§] francisco.alvarez-cuadrado@mcgill.ca.
[¶] francesco.amodio@mcgill.ca
[‖] markus.poschke@mcgill.ca

  - Code was last run in version 18.0 MP-Parallel Edition
  - Need packages:
    * `require`
    * `ftools`
    * `estout`
    * `winsor2`
    * `reghdfe`
  - The code was last run in a 2019 iMac, 3.6 GHz Quad-Core Intel Core i3 / 16 gb RAM, and took approximately 5 minutes.

- **Matlab**
  - Code was last run in version 2022b.
  - The code was last run on a 2021 Macbook Pro, Apple M1 Pro, and took approximately 3 minutes.
  - It requires some additional packages, which are all included in the `./Replication/model` folder.

5 Instructions for Replication

To create the dataset for the analysis starting from the raw LSMS-ISA data:

- In Stata, run the file `datasets-construction.do` located in the `./Replication/LSMS-ISA` folder. The only thing that the user needs to change is the main directory (to their particular directory) in line 13 of this file. The do-file installs some additional packages from Stata necessary to run the code.

To replicate the results in the paper:

- In Stata, run the file `paper-replication.do` located in the main `./Replication` folder. The only thing that the user needs to change is the main directory (to their particular directory) in line 13 of this file. The do-file installs some additional packages from Stata necessary to run the code.
- In MATLAB, run the file `generate_figures.m` located in the `./Replication/model` folder. This code reproduces the components of Figures 1, 2, 4 and 6.

All Tables and Figures are created and stored in `./Replication/output`.

6 List of Tables and Figures

All Tables and Figures are created and stored in `./Replication/output` with filenames `Table_XX` and `Figure_XX` and `XX` being their corresponding numbers in the paper.

6.1 List of Tables and Figures generated in Stata

This is the list of Tables and Figures files generated by the `paper-replication.do` located in the main `./Replication` folder.

Line Number	Output File	Extension	Line Number	Output File	Extension
396	descriptives/Table_1	xls	1821	Table_A16	tex
2371	descriptives/Table_3	xls	1858	Table_A17	tex
508	descriptives/Table_A1A	xls	996	Table_A18	tex
615	descriptives/Table_A1B	xls	1031	Table_A19	tex
721	descriptives/Table_A1C	xls	1113	Table_A2	tex
828	descriptives/Table_A1D	xls	2072	Table_A20	tex
2489	descriptives/Table_A32A	xls	2232	Table_A21	tex
2580	descriptives/Table_A32B	xls	2170	Table_A22	tex
2645	descriptives/Table_A32C	xls	2327	Table_A23	tex
2737	descriptives/Table_A32D	xls	2193	Table_A24	tex
2851	descriptives/Table_A33A	xls	2349	Table_A25	tex
2912	descriptives/Table_A33B	xls	2140	Table_A26	tex
1163	Figure_5	gph	2297	Table_A27	tex
1174	Figure_5	gph	2118	Table_A28	tex
1258	Figure_A1	gph	2275	Table_A29	tex
1269	Figure_A1	gph	1243	Table_A3	tex
1190	Figure_A2	gph	2095	Table_A30	tex
1202	Figure_A2	gph	2252	Table_A31	tex
1285	Figure_A2	gph	3073	Table_A34A	tex
1297	Figure_A2	gph	3097	Table_A34B	tex
1147	Table_2A	tex	3121	Table_A34C	tex
1963	Table_2B	tex	3145	Table_A34D	tex
3045	Table_4	tex	1687	Table_A35	tex
2052	Table_5	tex	1714	Table_A36	tex
1618	Table_A10	tex	1557	Table_A4	tex
1645	Table_A11	tex	1584	Table_A5	tex
1494	Table_A12	tex	1357	Table_A6	tex
1521	Table_A13	tex	1388	Table_A7	tex
1429	Table_A14	tex	1749	Table_A8	tex
1456	Table_A15	tex	1776	Table_A9	tex

6.2 List of Figures generated in Matlab

This is the list of Figures files generated by the file `generate_figures.m` located in the folder `./Replication/model`.

Line Number	Output File	Extension
174	Figure_1a_left	pdf
249	Figure_1a_middle	pdf
310	Figure_1a_right	pdf
172	Figure_1b_left	pdf
245	Figure_1b_middle	pdf
308	Figure_1b_right	pdf
174	Figure_2a_left	pdf
249	Figure_2a_middle	pdf
310	Figure_2a_right	pdf
172	Figure_2b_left	pdf
245	Figure_2b_middle	pdf
308	Figure_2b_right	pdf
359	Figure_3	pdf
427	Figure_4a_left	pdf
462	Figure_4a_middle	pdf
492	Figure_4a_right	pdf
547	Figure_4b_left	pdf
582	Figure_4b_middle	pdf
616	Figure_4b_right	pdf
689	Figure_6a	pdf
755	Figure_6b	pdf