

Replication files for “Information and Spillovers from Targeting Policy in Peru’s Anchoveta Fishery”

Gabriel Englander. American Economic Journal: Economic Policy.

DOI: <https://doi.org/10.5281/zenodo.7041706> (Note: This DOI is different from the DOI in the References section of the paper. That is because Zenodo only creates a DOI after I issue a release. The paper contains the final, publication-ready release DOI. The DOI I include here, 10.5281/zenodo.7041706, takes you to the newest release. So even though the two DOI are different, they take you to the same, final release version).

Data and Code Availability Statement

The paper uses public, non-confidential data from Peru’s Ministerio de la Producción (PRODUCE). The archive contains the data in the folder “Data/”. The files from PRODUCE are BE_2017to2019_allvessels.xlsx, landings_2017to2019.xlsx, closures_2014to2019.xlsx, and ownndf.csv. BE_2017to2019_allvessels.xlsx are the Bitácora Electrónica data for all vessels, landings_2017to2019.xlsx are the landings data for all vessels, closures_2014to2019.xlsx are the temporary spatial closures that PRODUCE declared between 2014 and 2019 (I only use those between 2017 and 2019 in the analysis), and ownndf.csv are ownership information and vessel characteristics for all vessels.

The paper uses public, non-confidential data from Sociedad Nacional de Pesquería (SNP). The archive contains the data in the folder “Data/”. The file from SNP is bedat_snp.csv. This file contains Bitácora Electrónica data for SNP vessels.

The paper uses public, non-confidential boundary data of Peru’s Exclusive Economic Zone from the Flanders Marine Institute (Flanders Marine Institute, 2012). The archive contains the data in the folder “Data/Intersect_IHO_EEZ_v2_2012”.

The paper uses public, non-confidential population-level length distribution data from Peru’s scientific agency (IMARPE, 2017). The archive contains the data in the folder “Data/”. The file from IMARPE is EvaluacionHidroacusticaRecPelagicosCrucero170304_march2017.jpg. This file is page 23 from their public report (IMARPE, 2017).

Computational Requirements

-Software: R. I used Version 4.1.2, but other versions should work too, especially those $\geq 4.1.0$.

-Packages: There are many of them. They are all recorded in renv.lock file. When you run Scripts/RUN_THIS_FIRST.R, the renv package will automatically install all of them.

-OS: I used Windows 10. Other versions of Windows, as well as Mac and Linux, should work too.

-CPU: I have Intel(R) Xeon(R) Gold 6132 CPU @ 2.60GHz 2.60 GHz (2 processors). This is the equivalent of 16 cores. If you have fewer than 16 cores, you will still be able to run the scripts, but your runtime will be longer.

-Memory: 128 GB. You can run most scripts with as little as 4 GB of memory though. A few scripts will require 128 GB of memory though, such as Scripts/make_figures/make_figureA11.R.

-Necessary disk space: 3 GB

-Wall clock-time: 92 hours. 64 hours for Scripts/make_data/7. make_data_figA13.R, 8.5 hours for Scripts/make_figures/make_figureA11.R, and 19.5 hours for all other scripts. If you don’t want to run Scripts/make_data/7. make_data_figA13.R, you may skip to make_figureA13.R since I provide the data necessary for creating Figure A13 in Output/Data/data_figA13.Rdata.

Instructions for Data Preparation and Analysis

Downloading and opening the replication files

If you are cloning the repository from Github (https://github.com/englander/replication_closures), open RStudio, click File -> New Project -> Version Control -> Git, paste “https://github.com/englander/replication_closures.git”, and click Create Project. If you downloaded the replication files as a zip file, extract them, open RStudio, click File -> Open Project, find `replication_closures.Rproj` among the files on your computer, and click Open.

Installing specific package versions

First, run `Scripts/RUN_THIS_FIRST.R`. That script will install all R packages you need. It installs the same package versions I used to facilitate reproducibility.

Data preparation

Run the scripts in `Scripts/make_data` folder in numeric order, starting with 0. `make_closures_df.R`. Scripts 0 to 6 take 5 hours to run in total. Script 7 takes 64 hours.

After running 4. `make_rddf.R`, you will have created the data necessary to create all tables, Figures 1-7, Figures A1-A11, Figures B1-B2, Figures C1-C2, and Figure E1. After running 5. `make_fleetthere_selfthere.R`, you will have created the data necessary to create Figures 8 and 9. Figure A12 requires running 6. `make_actualclosure_regressioncontrol.R`, and Figure A13 requires running 6. `make_actualclosure_regressioncontrol.R` and 7. `make_data_figA13.R`.

Note that 7. `make_data_figA13.R` requires 64 hours with 14 cores. The combined runtime of all other `make_data` scripts is five hours. I provide output from all `make_data` scripts in `Output/Data` folder. So if you want to skip the data preparation stage, and go right to reproducing tables and figures, you may do so. The only exception is Figure A12. Due to the size of the data used to create Figure A12, you must run 6. `make_actualclosure_regressioncontrol.R` on your computer to create `Output/TempData/actualclosure_regressioncontrol.Rdata`. Because of data provided in `Output/Data` folder, you could run 6. `make_actualclosure_regressioncontrol.R` without running the preceding `make_data` scripts.

Analysis

`Scripts/make_figures` folder contains the scripts that make all figures in the paper. Scripts are named by the figure(s) they create. The combined runtime for all `make_figures` scripts, excluding `make_figureA11.R`, is 5 hours. The runtime for `make_figureA11.R` is 8.5 hours.

`Scripts/make_tables` folder contains the scripts that make all tables in the paper. Scripts are named by the table(s) they create. The combined runtime for all `make_tables` scripts is 4 hours.

Files in `Scripts/other_empirics` folder contain calculations that are described in the paper but which do not produce a table or figure. The combined runtime for all `other_empirics` scripts is 5.5 hours.