Data and Code 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 owndf.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.xslx are the temporary spatial closures that PRODUCE declared between 2014 and 2019 (I only use those between 2017 and 2019 in the analysis), and owndf.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).

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.

Dataset list

Data file	Source	Notes	Provided
`Data/BE_2017to2019_allvessels.xlsx`	PRODUCE		Yes
`Data/landings_2017to2019.xlsx`	PRODUCE		Yes
`Data/closures_2014to2019.xlsx`	PRODUCE		Yes
`Data/owndf.csv`	PRODUCE		Yes
`Data/bedat_snp.csv`	SNP		Yes
`Data/Intersect_IHO_EEZ_v2_2012`	Flanders Marine Institute (2012)		Yes
`Data/EvaluacionHidroacusticaRecPelagicosCrucero170304_march2017.jpg`	IMARPE (2017)		Yes
`Output/Data/closed.Rdata`	Created by Scripts/make_data/0. make_closures_df.R		Yes
`Output/Data/matched_be_landings_level.Rdata`	Created by Scripts/make_data/1. match_be_landings.R		Yes

Data file	Source	Notes	Provided
Output/Data/grid2p.Rdata	Created by Scripts/make_data/1. match_be_landings.R		Yes
Output/Data/pbe_imp_uncorrected.Rdata	Created by Scripts/make_data/2. impute_size_be.R		Yes
Output/Data/pbe_imp.Rdata	Created by Scripts/make_data/3. correct_be.R		Yes
Output/Data/rddf_10km_lead1change_3dayrect.Rdata	Created by Scripts/make_data/4. make_rddf.R		Yes
Output/Data/fleetthere_selfthere.Rdata	Created by Scripts/make_data/5. make_fleetthere_selfthere.R		Yes
Output/TempData/prelim_data_figA13.Rdata	Created by Scripts/make_data/7. make_data_figA13.R		Yes
Output/Data/data_figA13.Rdata	Created by Scripts/make_data/7. make_data_figA13.R		Yes

Computational Requirements

Software Requirements

-Software: R. I used Version 4.1.2, but other versions should work too, especially those >= 4.1.0.

You may also need to install Rtools 4.0: https://cran.r-project.org/bin/windows/Rtools/rtools40.html

-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 CPUE @ 2.60GHz 2.60 GHz (2 processors). This is the equivalent of 16 cores. Some scripts hard-code parallel processing by specifying the number of cores to use. If you have fewer than 16 cores, calls like this will use all of your cores, and your runtime will be longer than the estimates provided here.

Controlled Randomness

-Random seed is set at line 71 of program Scripts/make_figures/make_figure6.R -Random seed is set at line 111 of program Scripts/make_figures/make_figureB1_figureB2.R

Memory and Runtime Requirements

-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: 86 hours. 64 hours for Scripts/make_data/7. make_data_figA13.R and 22 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.

Description of programs/code

-Scripts/make_data/0. make_closures_df.R cleans closures data and creates Output/Data/closed.Rdata.

-Scripts/make_data/1. match_be_landings.R matches Bitácora Electrónica data and landings data, creating matched_be_landings_belevel.Rdata.

-Scripts/make_data/2. impute_size_be.R constructs an uncorrected length distribution for all sets in the Bitácora Electrónica data, creating pbe_imp_uncorrected.Rdata.

-Scripts/make_data/3. correct_be.R corrects the Bitácora Electrónica data with the landings data, creating pbe_imp.Rdata.

-Scripts/make_data/4. make_rddf.R creates the potential closures data, which serves as input for many of the analysis scripts. The potential closures data file it creates is rddf_10km_lead1tolag4_3dayrect.Rdata.

-Scripts/make_data/5. make_fleetthere_selfthere.R creates the data necessary for creating Figures 8 and 9. The data file it creates is fleetthere_selfthere.Rdata.

-Scripts/make_data/6. make_actualclosure_regressioncontrol.R creates the data necessary for creating Figure A12. The data file it creates is actualclosure_regressioncontrol.Rdata.

-Scripts/make_data/7. make_data_figA13.R createsthe data necessary for creating Figure A13. The data file it creates is data_figA13.Rdata.

-Scripts/make_figures/... create the figure(s) referenced in the script file name. For example, Scripts/make_figures/make_figure1_figure4.R creates Figures 1 and 4.

-Scripts/make_tables/... create the table(s) referenced in the script file name. For example, Scripts/make_tables/make_table1.R creates Table 1.

-Scripts/other_empirics/appendix_C_robustness_length_distribution_imputation.R conducts a robustness check described in Appendix C. Instead of imputing length distribution of non-SNP sets at the two-week-of-sample by two-degree grid cell (as in the main specification of the paper), this script does so at the level of one-week-of-sample by one-degree-grid-cell. Then the script re-estimates the effect of the policy on juvenile catch.

-Scripts/other_empirics/appendix_D1.R estimates heterogeneous treatment effects by size of closure and length of closure period.

-Scripts/other_empirics/appendix_D2_firmsize.R estimates heterogeneous treatment effects by size of firm.

-Scripts/other_empirics/appendix_D2_vesselsize.R estimates heterogeneous treatment effects by vessel size.

-Scripts/other_empirics/appendix_D2_mediumvesselsonly.R estimates heterogeneous treatment effects by vessel size, among vessels owned by medium-sized firms.

-Scripts/other_empirics/discussion_alternative_policy.R simulates the effect on juvenile catch of replacing the closures policy with an alternative policy.

-Scripts/other_empirics/discussion_exports.R calculates the short-run effect of the policy on exports.

Instructions to Replicators

Some scripts hard-code parallel processing by specifying the number of cores to use. If you have fewer than 16 cores, calls like this will use all of your cores, and your runtime will be longer than the estimates provided here.

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 2.5 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.

List of tables and programs

Figure/Table #	Program	Line Number	Output file	Note
Figure 1	make_figures/make_figure1_figure4.R	227	figure1.pdf	
Figure 2	make_figures/make_figure2.R	153	figure2.pdf	
Figure 3	make_figures/make_figure3.R	208	figure3.pdf	
Figure 4	make_figures/make_figure1_figure4.R	284	figure4.pdf	
Figure 5	make_figures/make_figure5.R	108	figure5.pdf	
Figure 6	make_figures/make_figure6.R	232	figure6.pdf	
Table 1	make_tables/make_table1.R	482	table1.tex	
Table 2	make_tables/make_tableA1_table2.R	309	table2.tex	
Figure 7	make_figures/make_figure7.R	513	figure7.pdf	
Table 3	make_tables/make_table3.R	283	table3.tex	
Table 4	make_tables/make_table4.R	304	table4.tex	
Figure 8	make_figures/make_figure8_figure9.R	753	figure8.pdf	
Figure 9	make_figures/make_figure8_figure9.R	2038	figure9.pdf	

Figure/Table #	Program	Line Number	Output file	Note
Figure A1	make_figures/make_figureA1.R	504	figureA1.pdf	
Table A1	make_tables/make_tableA1_table2.R	259	tableA1.tex	
Figure A2	make_figures/make_figureA2_figureA3_figureA4_figureA5.R	232	figureA2.pdf	
Figure A3	make_figures/make_figureA2_figureA3_figureA4_figureA5.R	270	figureA3.pdf	
Figure A4	make_figures/make_figureA2_figureA3_figureA4_figureA5.R	306	figureA4.pdf	
Figure A5	make_figures/make_figureA2_figureA3_figureA4_figureA5.R	346	figureA5.pdf	
Figure A6	make_figures/make_figureA6.R	232	figureA6.pdf	
Figure A7	make_figures/make_figureA7.R	1667	figureA7.pdf	
Figure A8	make_figures/make_figureA8.R	1667	figureA8.pdf	
Figure A9	make_figures/make_figureA9.R	1786	figureA9.pdf	
Figure A10	make_figures/make_figureA10.R	473	figureA10.pdf	
Figure A11	make_figures/make_figureA11.R	1099	figureA11.pdf	
Figure A12	make_figures/make_figureA12.R	238	figureA12.pdf	
Figure A13	make_figures/make_figureA13.R	125	figureA13.pdf	
Figure B1	make_figures/make_figureB1_B2.R	179	figureB1.pdf	
Figure B2	make_figures/make_figureB1_B2.R	84	figureB2.pdf	
Figure C1	make_figures/make_figureC1.R	225	figureC1.pdf	
Figure C2	make_figures/make_figureC2.R	226	figureC2.pdf	
Table D1	make_tables/make_tableD1.R	15	tableD1.tex	
Table D2	make_tables/make_tableD2.R	105	tableD2.tex	
Figure E1	make_figures/make_figuree1.R	767	figuree1.pdf	

In-text numbers #	Program	Line Number	Output file	Note
5.7%	other_empirics/appendix_C_robustness_length_distribution_imputation.R	324	TempData/appendix_C_imputation_percent_change.dta	percent change in juvenile catch when I impute length distribution at one-week-of-sample by one-degree grid cell level
0.18	other_empirics/appendix_D1.R	520	TempData/appendix_D1_heterogeneity_area_level.dta	p-value on heterogeneous treatment effect by closure area
0.55	other_empirics/appendix_D1.R	779	TempData/appendix_D1_heterogeneity_days_level.dta	p-value on heterogeneous treatment effect by closure length
78%	other_empirics/appendix_D2_firm_size.R	403	TempData/appendix_D2_large_firm_fraction.dta	% of treatment effect from large-firm vessels
70%	other_empirics/appendix_D2_firm_size.R	414	TempData/appendix_D2_appendix_D2_juv_catch_by_firm_size.dta	% of juvenile catch by large-firm vessels, as well as fraction caught by medium-firm and singleton vessels
0.0005	other_empirics/appendix_D2_medium_vessels_only.R	407	TempData/appendix_D2_medium_vessels_only_reg.dta	p-value on heterogeneous treatment effect by vessel length, among medium-firm vessels only

In-text numbers #	Program	Line Number	Output file	Note
83%	other_empirics/appendix_D2_vessel_size.R	395	TempData/appendix_D2_above_median_length_vessel_juv_frac.Rdata	fraction of juvenile catch from above-median length vessels
-52%	other_empirics/discussion_alternative_policy.R	139	TempData/change_juv_catch_alternative_policy.Rdata	change in juvenile catch from replacing closures policy with an alternative
$75 million	other_empirics/discussion_exports.R	162	TempData/change_tons_exports.Rdata	change in exports due to policy

References

Flanders Marine Institute. 2012. "Intersect of IHO Sea Areas and Exclusive Economic Zones (version 2)." http://www.marineregions.org.

IMARPE. 2017. "Informe 'Evaluación Hidroacústica de Recursos Pelágicos' Crucero 1703-04." Instituto del Mar del Perú (IMARPE).