

Data and Code for “Does Patient Demand Contribute to the Overuse of Prescription Drugs?”

by Carolina Lopez, Anja Sautmann, and Simone Schaner (January 2021)

This document provides information needed to replicate the results in “Does Patient Demand Contribute to the Overuse of Prescription Drugs?”, Open ICPSR project ID openicpsr-126722.

The folders needed for the replication are the following:

- **Data**
- **Do_Files**
- **Tables**
- **Figures**
- **Survey instruments**

They should all be contained in the same folder (e.g. the folder they were archived in) in order to run the analysis .do files below.

Data

The top folder contains three data sets in Stata format.

1. LSS_analysis_datasets_2020_11_08.dta

The constructed data set used in the main analysis, collected in enumerator clinic visits (6 per clinic). Datasets were merged at the patient level.

1. Day-of-clinic records: clinic information filled out by the enumerator in the morning of the visit.

Questionnaire -> MorningCSCComID_Questionnaire (variables with suffix *_m*)

2. In-clinic patient survey: entry and exit interviews with patients at the clinic.

Questionnaire -> CSCComPatientInterview_Questionnaire

3. Home survey: home follow-up with subset of patients.

Questionnaire -> HomePatientInterview_Questionnaire (variables with suffix *_h*)

4. Intervention schedule: treatment assignment for each clinic-day (“official schedule”) (variables with suffix *_os*).

5. Internal records: administrative records on voucher redemption from intervention officers’ records on vouchers redeemed, including name of the patient which was used to match to the survey (variables with suffix *_pv* -patient voucher- and *_dv* -doctor voucher)

For the raw data, please visit the J-PAL DataVerse repository at <https://doi.org/10.7910/DVN/CLGN7Z>. To get access to these files, replicator must submit name, email, institution, position, and intended use of the files. Appendix "B.4 Variable Construction" describes how to create the variables used in the analysis.

2. LSS_CSCComCensus.dta

A census of the clinics included in the study, carried out in May 2016. Contains staffing and opening times.

3. LSS_DoctorEndlineInterview.dta

Endline interview with up to three providers (including head of the clinic) treating patients at the clinic that day, carried out after the intervention was concluded in December 2016/January 2017.

Requirements:

Stata version 13.1 or higher required.

Package needed for figures: grc1leg.pkg, coefplot.pkg

Packages needed for tables: pdsllasso, requires rlasso (st0594.pkg), command estadd and related commands from st0085_2.pkg

To setup these programs line 43 in “LSS_master_data_analysis.do” runs a program (“LSS_config_stata.do”) that installs all the packages needed (see below).

Computational environment: This package was validated on an Intel(R) Core(TM) i5-8250U CPU @ 1.60GHz 1.80GHz, RAM 8.00 GB, 64-Bit operating system. Stata/MP 14.0. To run the entire master file, it took approximately 3 hs. and 20 min. under this environment.

The guidelines below, sourced from <https://www.stata.com/products/compatible-operating-systems/>, outline computational requirements for Stata/MP and Stata/SE:

Platforms

Stata for Windows®

- Windows 10 *
- Windows 8 *
- Windows 7 *
- Windows Server 2019, 2016, 2012, 2008 R2 *

* Stata requires 64-bit Windows for x86-64 processors made by Intel® or AMD

Stata for Mac®

- Mac with Intel processor or Apple Silicon
- macOS 10.11 (El Capitan) or newer for Macs with Intel processors and macOS 11.0 (Big Sur) or newer for Macs with Apple Silicon

Stata for Linux

- Any 64-bit (x86-64 or compatible) running Linux

- For xstata, you need to have GTK 2.24 installed

Hardware requirements

Package	Memory	Disk space
Stata/MP	4 GB	1 GB
Stata/SE	2 GB	1 GB

Stata for Linux requires a video card that can display thousands of colors or more (16-bit or 24-bit color)

The creation of the bootstrap files can take a substantial amount of time, which will depend on your computational environment (e.g. 20 min. to several hours). Line 87 in “LSS_master_data_analysis.do” starts the bootstrap sampling; you will need to run this code before reproducing tables and figures that use bootstrapped standard errors.

Do Files:

Contains Stata .do files that replicate the data analysis (main paper and appendix).

- 1- “LSS_master_data_analysis.do”: This is in the top-level folder. To run, open this do file and set the path in line 22 to point to the folder on your computer that contains this archive. Running this .do file calls all the .do files needed to create the tables and figures in the paper. This file calls the programs below, which are organized into subfolders:
- 2- Subfolder “Stata admin files”: “LSS_config_stata.do”: In line 43, “LSS_master_data_analysis.do” opens a program that installs required packages in Stata, as was mentioned above.
- 3- Subfolder “Stata admin files”: “LSS_TablePrograms.do”: In line 48, “LSS_master_data_analysis.do” creates all the programs needed to create the tables.
- 4- Subfolder “Stata admin files”: “LSS_make_bootstrap_repsets.do”: In line 87, “LSS_master_data_analysis.do” creates all the files to run bootstrapped regressions.
- 5- Subfolder “Figures”: Contains codes to create all the figures of the paper and the appendix, these additional do files are called from “LSS_master_data_analysis.do”.
- 6- Subfolder “Tables”: Contains codes to create some tables of the paper and the appendix, these additional do files are called from “LSS_master_data_analysis.do”.

Tables: The .do files export tables in TeX format to the sub-folder *Tables*. Contains .tex files reproducing all the tables in the paper and appendix.

Table #	Program	Line Number in LSS_master_data_analysis.do
Table 1	1_table_control_group.do	156-159
Table 2	2_table_balance.do	163-175

Table 3	LSS_master_data_analysis.do	180-196
Table 4	LSS_master_data_analysis.do	200-219
Table 5	LSS_master_data_analysis.do	224-245
Table A1	not empirical	
Table B1	B1_table_clinic_census.do	254-256
Table B2	B2_table_means_doctor.do	260-262
Table B3	LSS_master_data_analysis.do	266-281
Table B4	LSS_master_data_analysis.do	285-297
Table B5	B5_table_selection.do	301-311
Table B6	B6_table_diff_test_consent.do	315-318
Table B7	B7_table_mal_prob_probit	322-325
Table B8	LSS_master_data_analysis.do	329-342
Table B9	LSS_master_data_analysis.do	346-365
Table B10	LSS_master_data_analysis.do	372-393
Table B11	LSS_master_data_analysis.do	397-424
Table B12	LSS_master_data_analysis.do	428-434
Table B13	LSS_master_data_analysis.do	438-449
Table B14	LSS_master_data_analysis.do	453-480
Table B15	LSS_master_data_analysis.do	485-504
Table B16	LSS_master_data_analysis.do	509-529

Figures: The .do files export figures in PDF format to the sub-folder *Figures*. The folder contains the figures produced by the .do files as well as two figures included in the paper not created as part of the data analysis (theory figure and intervention schedule).

Figure #	Program	Line Number in LSS_master_data_analysis.do
Figure 1	not empirical	
Figure 2	not empirical	
Figure 3	3_graph_outcomeXpredpos_control.do	95-98
Figure 4	4_graph_tmtXpredpos.do	102-105
Figure 5	5_graph_voucher_tmt.do	109-112
Figure B1	B1_graph_misallocation.do	121-124
Figure B2	B2_graph_dist_pred_pos.do	128-131
Figure B3	B3_graph_outcomeXpredpos_control_home.do	135-138
Figure B4	B4_graph_vouchersXbins.do	142-145

Discrepancies: The code and data files provided should enable the analyst to exactly replicate the results in the paper, with the (potential) exception of the figures that use Stata’s “lpoly” command (Figures 3, 4 and B3). When validating the replication, the authors found that the graphs render slightly differently when run using Stata 16 versus 14 software. This issue persists even though the Stata version is set to 13.1 at the beginning of the master do file. The authors believe this issue is caused by differences in Stata’s releases. The graphs in the manuscript were generated using Stata 14.2 software.

Survey Instruments: contains all survey instruments used for primary data collection. Note: the census data is in French.

To make it easy for the user to compile all the files, we have added LSS_Main_Replication.tex and LSS_Appendix_Replication.tex which include all the figures and tables of the analysis, as well as the pdf output from these files.

Data and Code References

Lopez, Carolina, Anja Sautmann, and Simone Schaner. 2021. "Replication data for: Does Patient Demand Contribute to the Overuse of Prescription. Drugs?" American Economic Association [publisher], Inter-university Consortium. for Political and Social Research [distributor].

Sautmann, Anja, and Simone Schaner. 2017. "Incentives for Accurate Diagnosis: Improving Health Care Quality in Mali." Harvard Dataverse, V1, <https://doi.org/10.7910/DVN/CLGN7Z>.