Readme File for the Replication of

"Path Dependency in Physician Decisions"

Data Availability Statement.

For the analysis of physicians' decisions in the emergency department (ED) setting, we use hospital administrative data obtained from a proprietary source in Southeast Asia (SAAH, 2013). To strictly observe the confidentiality of the data, as per specified by the non-disclosure agreement, we would not be able to post the data on public domain. Nonetheless, we understand that ensuring replicability is essential to any scientific endeavour. We extract variables used to replicate all the results reported in the paper and the online appendix (EDSample.dta in the package), and publish all computer code used for the analysis. We have submitted all original data and computing codes to a Research Data Repository (RDR) located at the National School of Department (NSD), Peking University. The RDR is to safeguard researchers' academic integrity and is immune to subsequent data alteration after submission. Interested researchers could refer to michaelshen@nsd.pku.edu.cn for further information on how to access the original data.

For the analysis of physicians' decisions for childbirth, we use the New York hospital inpatient data (NYSDOH, 2015). The data are confidential and were obtained under Data Use Agreements with the New York State Department of Health Statewide Planning and Research Cooperative System (SPARCS). They cannot be shared publicly. Researchers interested in accessing the data may contact SPARCS representatives at sparcs.requests@health.ny.gov. For more details, see https://www.health.ny.gov/statistics/sparcs/access/.

Computational Requirements.

Software Requirements

- Stata/MP (code was last run with version 16.0)
  - outreg2 (as of 2014-08-17)
  - fsum (as of 2014-05-06)

    - grc1leg2 (as of 2023-05-14)
    - the program "0_setup.do" will install all dependencies locally, and should be run once prior to running other programs.
  - Matlab (code was run with Matlab Release 2023a)

Memory and Runtime Requirements

The code in the ED folder was last run on a laptop 14-core i9-12900H CPU @ 2.50GHz with 32 GB of RAM. One run of the Matlab code took around 3.5 hours. One run of the Stata code took around 10 minutes for the tables and 40 minutes for the figures.

The code in the childbirth folder was last run on a Cornell Restricted Access Data Center (CRADC) server workstation 18-core Intel Xeon Gold 6154 CPU @ 3.00GHz with 225 GB of RAM. One run of the data cleaning code took about 3 hours. One run of the analysis code took about 8 hours.

Included Files

ED folder

This folder includes the data and all the programs for the analysis of physicians' decisions in the ED setting.

- Stata folder
  - Data file

    EDSample.dta (data used to replicate all the results for the ED setting reported in the paper and the online appendix)

  - Code files
    - 0_setup.do (install needed packages)
    - master.do (run all the do files)
    - main tables.do (replicate Tables 1–7 and Table 9 in the paper)
    - appendix tables.do (replicate Tables A.1–A.7 in the online appendix)

- Figure 1&A3.do (reproduce Figure 1 in the paper and Figure A.3 in the online appendix)
- Figure A2.do (reproduce Figure A.2 in the online appendix)
- Figure A4.do (reproduce Figure A.4 in the online appendix)
- placebo.ado (placebo program used in Figure A4.do)

– Instructions for replicators
  - In Stata, run programs "0_setup.do" once on a new system to set up the working environment.
  - Run "master.do" to run all steps in sequence.

- Matlab folder
  – Data file
    dataMatlab.csv (data used to reproduce Table 10 in the paper)
  – Code files
    - Main_Table10.m (base)
    - getopt.m (function)
    - ll.m (function)
  – Instructions for replicators
    - In Matlab, run "Main_Table10.m" to reproduce Table 10 in the paper.

<u>Childbirth folder</u>

This folder includes all the programs for the analysis of physicians' decisions in the childbirth setting.

- Data cleaning codes: ny_clean.do

  The raw datasets contain rich information on all hospital inpatient admissions that occurred in the State of New York between 2005-2015.

  Childbirths were identified by searching across all diagnosis and procedure variables for the following ICD-9 codes:

- ICD-9 Diagnosis codes: 630-676.9, V22, V23
- ICD-9 Procedure codes: 72-75

If any of the above ICD-9 code was recorded in any of the diagnosis or procedure variables in the dataset, the admission was identified as a childbirth delivery.

All childbirth observations were saved as "ny_data.dta" and then cleaned using the attached "ny_clean.do" file.

- Data analysis codes: ny_analysis.do
  - Reproduce Table 8 in the paper and appendix Figure B1, Table B1, and Table B2.

Data Citations

Southeast Asian Anonymous Hospital (SAAH). 2013. "Emergency Department Administrative Data, 2011–2012." Unpublished Data. Accessed January 29, 2018.

New York State Department of Health (NYSDOH). 2015. "New York Statewide Planning and Research Cooperative System (SPARCS) Dataset, 2005–2015". Available at: https://www.health.ny.gov/statistics/sparcs/access/. Accessed June 1, 2020.