OVERVIEW
==========
Replication code for

Long-Term Care Hospitals: A Case Study in Waste

by Einav, Finkelstein, and Mahoney

Production date: February 11, 2021

DATA ACCESS: RESTRICTED DATA
==========
The code in this archive requires access to Medicare claims data from the Centers for Medicare and Medicaid Services (CMS). The code is designed to access these data via a Subversion repository we designed for this project. Researchers wishing to replicate our results will need to obtain their own copy of the data via an independent Data Use Agreement with CMS. The process for obtaining a DUA can be found through:
https://www.cms.gov/Research-Statistics-Data-and-Systems/Files-for-Order/Data-Disclosures-Data-Agreements/Overview
The code will need to be updated to reflect the format of the researcher's own copy of the data and to point to the appropriate repository or storage device.

The analysis also requires data from other sources that are not publicly available. In particular, data from the American Hospital Association were used to measure hospital characteristics. We use AHA data stored on NBER servers. Documentation and information on accessing AHA data through the NBER can be found at:
https://www.nber.org/research/data/annual-survey-hospitals

DATA ACCESS: PUBLIC DATA
==========
The paper also draws on a number of publicly available data sources. These are stored under /code/raw/. The links and descriptions can be found in the readmes located in the relevant subdirectories.

We do not provide data for the Provider of Services Files under /code/raw/POS 1984-2016/. As stated in the readme of this subdirectory, the Provider of Services Files can be downloaded from this NBER website:
https://www.nber.org/research/data/provider-services-files

HOW TO USE
==========
Our code directory is divided into a handful of main directories:

- directories in /raw/ store raw data
- directories in /derived/ take data as inputs (often from /raw/) and produce data as outputs
- directories in /derived_local/ also take data as inputs (often from /raw/) and produce data as outputs
- directories in /analysis/ take data as inputs (from /derived/ & /derived_local/)

  and produce results as outputs
- directories in /lib/ store code libraries that are loaded and used by other
  directories

Here we describe the contents of each code directory such as those in /derived/, /derived_local/, and /analysis/. Each directory is self-contained so that the code can be run without inputs from outside the directory. This is accomplished by constructing symbolic links or version-controlled references to datasets or files necessary for analysis.

Most directories have a similar structure:

/code/
    Contains all scripts and code necessary to execute the directory. In most cases there will be a Stata do file or a SAS script with macros that perform the main functionality.

/code/externals.txt
    Lists the external resources needed to run the directory, calling versions of data or programs saved in one of the Subversion repositories used for the project. Items referenced in this file will be populated in the /externals/ folder in directory. This file can be used to locate other portions of code that create datasets used in an analysis.

/code/links.txt
    References datasets that are too large for version control and so are stored locally on the machine. Items referenced in this file will be populated as symbolic links under /external_links/ in the directory. This can also be used to located other portions of code that create datasets used in an analysis.

/code/make.py
    Executes all of the code in the directory. You will not be able to run this code, but you can use it as documentation for the correct order in which the other scripts are executed.

/externals/
    Contains data not created at runtime that are necessary to execute the directory. The file /code/externals.txt lists its contents.

/external_links/
    Contains symbolic links to data not created at runtime that are necessary to execute the directory. The file /code/links.txt lists its contents.

/output/
    Contains log files, txt or Excel files with tables, eps files with figures, csv files with data, or dta files with Stata datasets. This is populated once the /code/ directory is executed.

In this replication archive, the directories /lib/, /raw/, /analysis/, /derived/, and /derived_local/ live under the /code/ directory. Within these directories, we

organized our subdirectories according to the path files found in externals.txt and links.txt files. That is, each line of the txt files will show specific path files such as "..../trunk/derived/LTCH/.....". These path files match up with the different subdirectories we have made.

BUILDING FROM THE RAW DATA
==========
The project is built from a series of raw Medicare files. These data are stored on the NBER servers and were created by Jean Roth at the NBER. These include:

/code/raw/MedPAR (100 pct)/
    Contains symbolic links to 100 percent MedPAR files at the NBER.

/code/raw/Medicare Centralized Denom (100 pct)/
    Contains symbolic links to the 1998 - 2014 100% Medicare Denominator (BSF) extracts at the NBER.

/code/raw/HHA (100 pct)/
    Contains symbolic links to the 2001 - 2014 100% Medicare Home Health Agency Claims at the NBER.

/code/raw/Hospice (100 pct)/
    Contains symbolic links to the 2002 - 2014 100% Medicare Hospice Claims at the NBER.

/code/raw/Cost and Use (100 pct)/
    Contains symbolic links to the 1999 - 2014 100% Medicare Cost and Utilization Segment files of the Master Beneficiary Summary file at the NBER.

Subdirectories in /code/derived/LTCH/ and /code/derived_local/LTCH/ build off these raw data to create patient outcomes and panel datasets. These outputs feed into /code/derived_local/LTCH/Patient Level Event Study/, which contains the primary analysis sample we use for the event studies.

TABLE CROSSWALK
==========

Table	Code Location	Text File Name
1	/analysis/LTCH/Covariates by LTCH Existence/	covariates.txt
2	/analysis/LTCH/Sum Stats High Phat/	covs_sum_stats.txt
3	/analysis/LTCH/Sum Stats High Phat/	sum_stats.txt
4	/analysis/LTCH/Make ES Tables/	summary_table.txt
5	/analysis/LTCH/Make ES Tables/	het_table.txt

Appendix	Code Location	Text File Name
A.1	/analysis/LTCH/Sum Stats High Phat/	sum_bins.txt
A.2	/analysis/LTCH/Make ES Tables/	iv_ests_by_cell.txt
A.3	/analysis/LTCH/Make ES Tables/	mort.txt
A.4	/analysis/LTCH/Make ES Tables/	recs.txt
A.5	/analysis/LTCH/Make ES Tables/	pre_share.txt

A.6	/derived_local/LTCH/Predict LTCH Discharged No CCs/	main_tree.pdf
A.7	/analysis/LTCH/Phat Distribution/	distr_alldists.txt
A.8	/analysis/LTCH/Make ES Tables/	alt_middle.txt
A.9	/analysis/LTCH/Make ES Tables/	robust.txt
A.10	/analysis/LTCH/Make ES Tables/	compare_pooled_combined.txt
A.11	/analysis/LTCH/Sum Stats High Phat/	ndays_sum_stats.txt
A.12	/analysis/LTCH/Make ES Tables/	ndays_compare.txt
A.13	/analysis/LTCH/Make ES Tables/	spend.txt

FIGURE CROSSWALK

Figure	Code Location	Figure Name
1	/analysis/LTCH/Provider Maps Over Time/	ltch_over_time.pdf
2a	/analysis/LTCH/Provider Maps Over Time/	hsa_color_1984.pdf
2b	/analysis/LTCH/Provider Maps Over Time/	hsa_color_1998.pdf
2c	/analysis/LTCH/Provider Maps Over Time/	hsa_color_2014.pdf
3	/analysis/LTCH/Providers Panel Structure Quarterly/	hist_enter_quarter.pdf
4	/analysis/LTCH/Make ES Figures/	rpart_fit_bin0.pdf
5	/analysis/LTCH/Make ES Figures/	DestLTCH_bin5.pdf
6a	/analysis/LTCH/Make ES Figures/	DestSNFIRF_bin5.pdf
6b	/analysis/LTCH/Make ES Figures/	DestOther_bin5.pdf
6c	/analysis/LTCH/Make ES Figures/	DestBabyLTCH_bin5.pdf
6d	/analysis/LTCH/Make ES Figures/	DestInpat_bin5.pdf
6e	/analysis/LTCH/Make ES Figures/	DestDeath_bin5.pdf
7a	/analysis/LTCH/Make ES Figures/	cur_post_ach_util_day_bin5.pdf
7b	/analysis/LTCH/Make ES Figures/	cur_post_ach_mdcr_spend_bin5.pdf
8a	/analysis/LTCH/Make ES Figures/	cur_post_ach_oop_bin5.pdf
8b	/analysis/LTCH/Make ES Figures/	home_90days_bin5.pdf
8c	/analysis/LTCH/Make ES Figures/	mort90day_bin5.pdf

Appendix	Code Location	Figure Name
A.1	/analysis/LTCH/Make ES Figures/	DestLTCH_bin0.pdf
A.2a	/analysis/LTCH/Make ES Figures/	DestSNFIRF_bin0.pdf
A.2b	/analysis/LTCH/Make ES Figures/	DestOther_bin0.pdf
A.2c	/analysis/LTCH/Make ES Figures/	DestBabyLTCH_bin0.pdf
A.2d	/analysis/LTCH/Make ES Figures/	DestInpat_bin0.pdf
A.2e	/analysis/LTCH/Make ES Figures/	DestDeath_bin0.pdf
A.3a	/analysis/LTCH/Make ES Figures/	cur_post_ach_util_day_bin0.pdf
A.3b	/analysis/LTCH/Make ES Figures/	cur_post_ach_mdcr_spend_bin0.pdf
A.4a	/analysis/LTCH/Make ES Figures/	cur_post_ach_oop_bin0.pdf
A.4b	/analysis/LTCH/Make ES Figures/	home_90days_bin0.pdf
A.4c	/analysis/LTCH/Make ES Figures/	mort90day_bin0.pdf
A.5a	/analysis/LTCH/Make ES Figures/	post_ach_ltch_days_bin5.pdf
A.5b	/analysis/LTCH/Make ES Figures/	post_ach_ltch_mdcr_spend_bin5.pdf
A.5c	/analysis/LTCH/Make ES Figures/	post_ach_snf_irf_days_bin5.pdf
A.5d	/analysis/LTCH/Make ES Figures/	post_ach_snf_irf_mdcr_spend_bin5.pdf
A.5e	/analysis/LTCH/Make ES Figures/	util_day_bin5.pdf

A.5f	/analysis/LTCH/Make ES Figures/	mdcr_spend_bin5.pdf
A.6a	/analysis/LTCH/Make ES Figures/	DestLTCH_over3.pdf
A.6b	/analysis/LTCH/Make ES Figures/	DestLTCH_not_over3.pdf
A.6c	/analysis/LTCH/Make ES Figures/	DestLTCH_over8.pdf
A.6d	/analysis/LTCH/Make ES Figures/	DestLTCH_not_over8.pdf
A.6e	/analysis/LTCH/Make ES Figures/	DestLTCH_ventilation.pdf
A.6f	/analysis/LTCH/Make ES Figures/	DestLTCH_not_vent.pdf
A.7a	/analysis/LTCH/Make ES Figures/	DestLTCH_half_1.pdf
A.7b	/analysis/LTCH/Make ES Figures/	DestLTCH_half_2.pdf
A.7c	/analysis/LTCH/Make ES Figures/	DestLTCH_q_1.pdf
A.7d	/analysis/LTCH/Make ES Figures/	DestLTCH_q_4.pdf
A.7e	/analysis/LTCH/Make ES Figures/	DestLTCH_forprof.pdf
A.7f	/analysis/LTCH/Make ES Figures/	DestLTCH_not_forprof.pdf
A.8	/analysis/LTCH/Analyze Claim Coded Events/	scatter_enter_24_wkly.pdf
A.9	/analysis/LTCH/Analyze Claim Coded Events/	claims_POS_yr_enter.pdf
A.10	/analysis/LTCH/Analyze Baby LTCHs/	time_exist_hist.pdf
A.11	/analysis/LTCH/Make ES Figures/	DestInpat2_bin5.pdf
A.12	/derived_local/LTCH/Predict LTCH Discharges No CCs/	tuning.pdf
A.13a	/analysis/LTCH/Make ES Figures/	tracheostomy_bin5.pdf
A.13b	/analysis/LTCH/Make ES Figures/	ventilation_bin5.pdf
A.14a	/analysis/LTCH/Make ES Figures/	DestLTCH_middle2.pdf
A.14b	/analysis/LTCH/Make ES Figures/	DestLTCH_middle3.pdf
A.15a	/analysis/LTCH/Make ES Figures/	DestLTCH_bin5.pdf
A.15b	/analysis/LTCH/Make ES Figures/	DestLTCH_control_existing.pdf
A.15c	/analysis/LTCH/Make ES Figures/	DestLTCH_balanced.pdf
A.15d	/analysis/LTCH/Make ES Figures/	DestLTCH_all.pdf
A.15e	/analysis/LTCH/Make ES Figures/	DestLTCH_no_mcaid.pdf
A.15f	/analysis/LTCH/Make ES Figures/	DestLTCH_cnty.pdf
A.16a	/analysis/LTCH/Make ES Figures/	cur_post_ach_mdcr_spend_bin5.pdf
A.16b	/analysis/LTCH/Make ES Figures/	cur_post_ach_outlramt_bin5.pdf
A.16c	/analysis/LTCH/Make ES Figures/	cur_post_ach_mdcr_plus_bin5.pdf

CODE LIBRARIES

We provide some of the Stata ado files and other code modules necessary to replicate our results. These scripts can be found under /code/lib/. This also includes SAS programs.

The file /code/lib/input_parameters/ contains files with lists of Stata global macros and their corresponding values that can be loaded into Stata.