Countervailing Market Power and Hospital Competition

Eric Barrette, Gautam Gowrisankaran, and Robert Town

Readme file for understanding code

December 8, 2020

We include the SQL code in .txt files to extract data from the Health Care Cost Institute (HCCI) database and the Stata code in .do files to construct the analytic dataset and perform analyses. We performed all data processing and analysis in the HCCI Enclave using a Vertica SQL server (for the SQL code) and Stata 14.1 (for the Stata code).

We ran the code with the following steps:

1) We ran SQL data extractions using the SQL script "IP claims pull from database.txt."
2) We ran the Stata data processing code by running the Stata file "00 Data.do."
3) We ran the Stata data analysis code by running the Stata file "00 Analysis.do."

We include the SQL code and Stata data processing code in the "Data" subfolder and the analysis code in the "Analysis" subfolder. Within the HCCI environment, we stored the code in a "Scripts" subfolder within the Data or Analysis subfolders along with "Input," "Temp," and "Output" subfolders. These folders included input data sets, temporary files, and data extracts/results, respectively. We defined the paths with global macros in the.do files that we run, "00 Data.do" and "00 Analysis.do."

Our main source of data is administrative claims from HCCI. The HCCI data include claims from beneficiaries with employer-sponsored coverage from Aetna, Humana, and UnitedHealthcare. We include the HCCI data dictionary, HCCI_Data_Dictionary_7.17.18, in the data subfolder. This dictionary contains details of the included fields. For information about obtaining these claims data for replication, contact HCCI: https://healthcostinstitute.org/.

As detailed in the paper, we use four other sources of data:

1) We use Medicare DRG weights for the hospital price severity adjustment. They are available in the Inpatient Prospective Payment System Final Rule files at: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/IPPS-Regulations-and-Notices.
2) We use the American Hospital Association (AHA) Annual Survey database. For information about obtaining these data, contact the AHA: https://www.ahadata.com/. The HCCI data use agreement prohibits linking the AHA data directly to the HCCI data because the AHA data includes provider identifying information (e.g., hospital name or address). Using an AHA license, we link the HCCI database with a deidentified version of the AHA database within the HCCI environment. Our data includes a masked NPI—denote HNPI—that we reference throughout the code.
3) We use American Medical Association (AMA) annual competition reports, available at: https://commerce.ama-assn.org/store/ui. The insurer HHI used in Step 3 of the analysis

derive from these data. We collected annual, CBSA, and plan type (POS and PPO) specific HHIs for the years 2011-2014.

4) We use socioeconomic data from the U.S. Census data as controls in Step 3 of the analysis. Uninsurance rate data are available at: https://www.census.gov/programs-surveys/sahie.html. Household income data are available at: https://www.census.gov/programs-surveys/saipe.html.