Problem Set 6

Milton Straw
ECON 815: Topics in Microeconomics

Fall 2019

PART A: Data

For this assignment, I explored the websites for the Office of Disease Prevention and Health Promotion (ODPHP) and the Center for Medicare and Medicaid Services (CMS). After looking through several datasets, I decided on some data from the CMS regarding the Inpatient Prospective Payment System (IPPS). I selected this dataset for the availability of an API, its tractability, and its relevance to my research interests. Here is the Data.CMS.gov data that I used as the base for this assignment.

This dataset contains information regarding hospital charges and Medicare repayments for 1,000 hospitals across the United States. Each of the charges is labeled by a DRG, so that like services can be compared. For the sake of this exercise, only one service DRG is used, resulting in 1,000 observations. Here are some brief explanations of the less obvious variables...

- average_covered_charges: The provider's average charge for services covered by Medicare for all discharges in the DRG.
- average_medicare_payments: The average of Medicare payments to the provider for the DRG including the DRG amount, teaching, disproportionate share, capital, and outlier payments for all cases. Also included are co-payment and deductible amounts that the patient is responsible for.
- drg_definition: DRGs are a classification system that groups similar clinical conditions (diagnoses) and the procedures furnished by the hospital during the stay.
- total_discharges: The number of discharges billed by the provider for inpatient hospital services.

To retrieve the data from the CMS, I used the "Request" and "json" packages for Python. Using the provided .json formatted api, I read the .json text into a dataframe using Pandas. From here, I was able to organize the data to my liking, and perform the manipulations necessary for my analysis.

PART B: Analysis

Using the data I retrieved, I attempted to gain some insight to the question: ”How are hospital charges and Medicare repayments for a given procedure affected by the number of hospitals in a given region?”. My intuition behind this simple questions is that a greater density of hospitals might be a proxy for competition, and greater competition would drive charges downward. It could, however, be the case that greater competition within a region gives Medicare more bargaining power for repayments.

Due to the time constraints associated with a problem set, I decided to seek visualizations that might provide insight into my question before designing a model. In Figure 1, I simply cluster observations of Average Covered Charges (amount billed) by the number of hospitals within the region for that observation. Figure 2 does that same for Percentage of Covered Charges Paid by Medicare. This was found by performing a simple calculation of Average Covered Charges divided by Average Medicare Payments. These cluster visualizations were chosen so that I could see every observation since there aren't too many of them, and decipher their density.

Figure 1: Billable Amounts

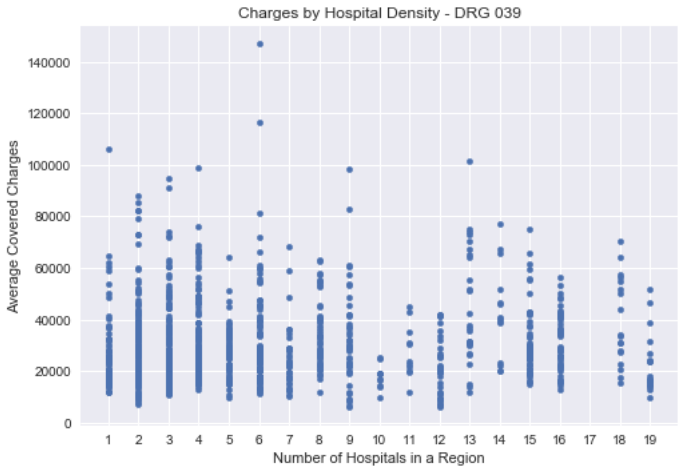
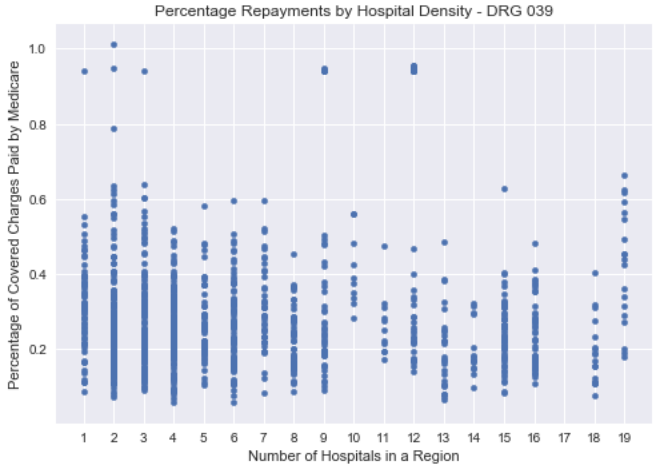


Figure 2: Medicare Repayment



I believe that a pattern is visible in these figures, but it is slight. Replicating the visualizations with the full 163,000 observations available in the data might suggest that some exploratory analysis is warranted. The suggestion here, and in Figures 3 & 4, show that Average Covered Charges appears to increase as the number of hospitals in a region increases; and that Percentage of Covered Charges Paid by Medicare appears to decrease as the number of hospitals in a region increases.

Figure 3: Linear Fit - Billable Amounts

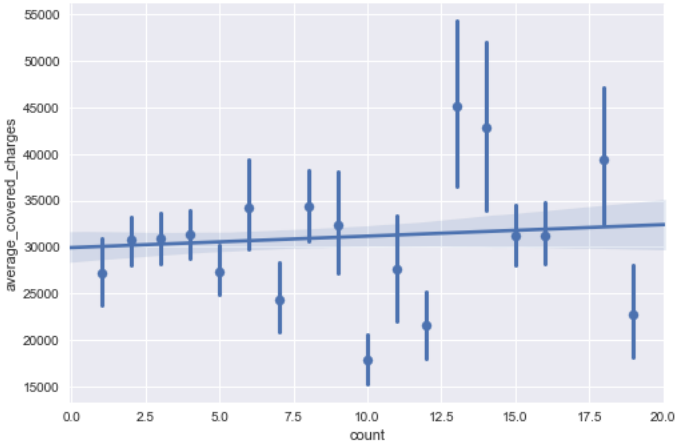
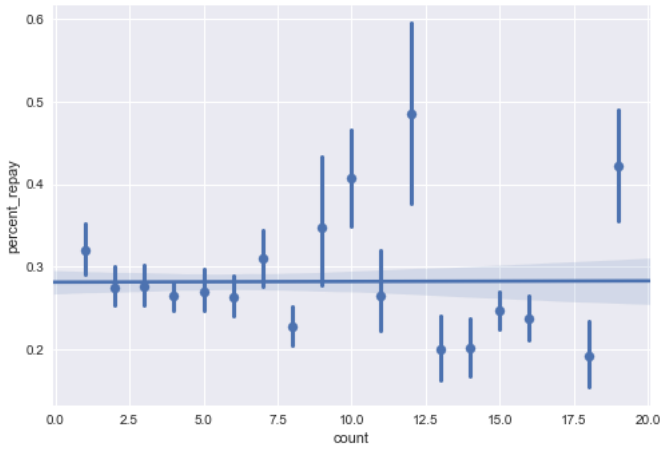


Figure 4: Linear Fit - Medicare Repayment



This suggests that hospitals charge more for a given procedure in an area with greater hospital density (more competition), but Medicare repays a smaller percentage of those charges.