Homework 4

Instructions

In this assignment, you'll turn to the Hospital Cost Report Information System data. These data are described in detail in the HCRIS GitHub Repo. The due date for initial submission is **3/31**, the revision due date is **4/2**, and the final due date is Friday, **4/3**.

Summarize the data

1. From 2009-2019, how many hospitals filed more than one report in the same year? Show your answer as a line graph of the number of hospitals over time.

2. After removing/combining multiple reports, how many unique hospital IDs (Medicare provider numbers) exist in the data?

3. From 2009-2019, what is the distribution of total charges (tot_charges in the data) in each year? Show your results with a "violin" plot, with charges on the y-axis and years on the x-axis. For a nice tutorial on violin plots, look at Violin Plots with ggplot2.

4. From 2009-2019, what is the distribution of estimated prices in each year? Again present your results with a violin plot, and recall our formula for estimating prices from class. Be sure to do something about outliers and/or negative prices in the data.

```
discount_factor = 1-tot_discounts/tot_charges
price_num = (ip_charges + icu_charges + ancillary_charges)*discount_factor - tot_mcare_paymer
price_denom = tot_discharges - mcare_discharges
price = price_num/price_denom
```

5. What share of hospitals are penalized under the HRRP/VBP? Provide a graph showing the share of penalized hospitals over time, from 2012-2019.

Estimate ATEs

Now let's work on implementing an instrumental variables estimator. Specifically, we're going to estimate the effects of HRRP/VBP penalties on hospital prices, using prior Medicare discharges (e.g., exposure to the program) as an instrument. When explaining your findings, try to limit your discussion just to a couple of sentences. As we did in class, for this we're going to focus only on a cross-sectional analysis of price changes from 2011 to 2014 as a function of 2012 penalties and average pre-2012 Medicare discharges (2009-2011). Define penalty as the net penalty under both HRRP and VBP.

6. Provide a summary of OLS estimates of the effect of net penalties on price changes. Present your results in a table with three different specifications: 1) a "baseline" specification using only net penalty as a covariate; 2) "baseline" specification plus the pre-penalty (2009-2011) mean bed size; 3) "baseline" specification plus bed size plus pre-penalty (2009-2011) average Medicaid discharges.

7. Provide a scatterplot of net penalty against pre-2012 Medicare discharges.

8. Provide a summary of the first stage and reduced-form results using pre-penalty Medicare discharges as an instrument for net penalties. Present your results in a table with three different specifications as in Question 6.

9. Provide a summary of IV estimates of the effect of net penalties on price changes. Again present your results in a table with the three different specifications as in Questions 6 and 8.

10. Briefly explain the "Local" ATE in the context of your estimates. How might a local effect differ from an overall ATE in this setting?