Homework 3

Instructions

In this assignment, you'll again work with the Medicare Advantage GitHub repository. This time you'll also need to incorporate the star ratings information, and we'll focus on data from 2010 through 2015.

The due date for initial submission is **3/3**, the revision due date is **3/5**, and the final due date is Friday, **3/6**.

Summarize the data

1. Provide a table of summary statistics showing the mean star rating, mean number of enrollments, and mean market share for plans by year. The variables (star rating, enrollments, market share) should be listed as columns and years (2010-2015) as rows. Provide an additional column reflecting the total number of plans underlying your summary statistics in each year.
2. Repeat part 1 but focusing only on plans without a star rating. Naturally, in this case, you need only present the mean enrollments and market share, not the mean star rating, along with a column showing the count of all such plans in each year.
3. Provide bar graphs showing the distribution of star ratings in 2010, 2012, and 2015. How has this distribution changed over time?
4. Provide a table showing the regression resuls from an ordinary least squares regression of market share on star ratings, again for each year from 2010 through 2015. In this table, the rows should reflect your coefficient estimates and the columns should reflect different estimates for each year. In your regression specifications, please treat star ratings of 2.5 or below as your excluded category, and include indicator variables for star ratings of 3, 3.5, 4, and 4.5 or above.

Estimate ATEs

For the rest of the assignment, we'll use a regression discontinuity design to estimate the average treatment effect from receiving a marginally higher rating. We'll focus only on 2010.

5. Calculate the running variable underlying the star rating. Provide a table showing the number of plans that are rounded up into a 3-star, 3.5-star, 4-star, 4.5-star, and 5-star rating.
6. Using the RD estimator with a bandwidth of 0.125, provide an estimate of the effect of receiving a 3-star versus a 2.5 star rating on enrollments. Repeat the exercise to estimate the effects at 3.5 stars, and summarize your results in a table.
7. Repeat your results for bandwidhts of 0.1, 0.12, 0.13, 0.14, and 0.15 (again for 3 and 3.5 stars). Show all of the results in a graph. How sensitive are your findings to the choice of bandwidth?
8. Examine (graphically) whether contracts appear to manipulate the running variable. In other words, look at the distribution of the running variable before and after the relevent threshold values. What do you find?
9. Examine whether plans just above the threshold values have different characteristics than contracts just below the threshold values. Use HMO and Part D status as your plan characteristics.
10. Summarize your findings from 5-9. What is the effect of increasing a star rating on enrollments? Briefly explain your results.