# ECON 470 HW4

Ben Yang

2023-04-03

## 1 Instructions

In this assignment, you'll again work with the Medicare Advantage data. These data are described in detail in the Medicare Advantage GitHub Repo. We worked with a subset of these data back in assignment 1; however, this assignment requires that you work with a more complete version of the Medicare Advantage data, but for fewer years. Once you have the data downloaded and the code running, answer the following questions:

## 2 Summarize the data

#### 2.0.1 Question 1. Remove all SNPs, 800-series plans, and prescription drug only plans (i.e., plans that do not offer Part C benefits). Provide a box and whisker plot showing the distribution of plan counts by county over time. Do you think that the number of plans is sufficient, too few, or too many?

Figure 1 displays the distribution of plan counts by county over time from 2007 to 2015. The box and whisker plot shows that there are many extreme values on the upside of the distribution.

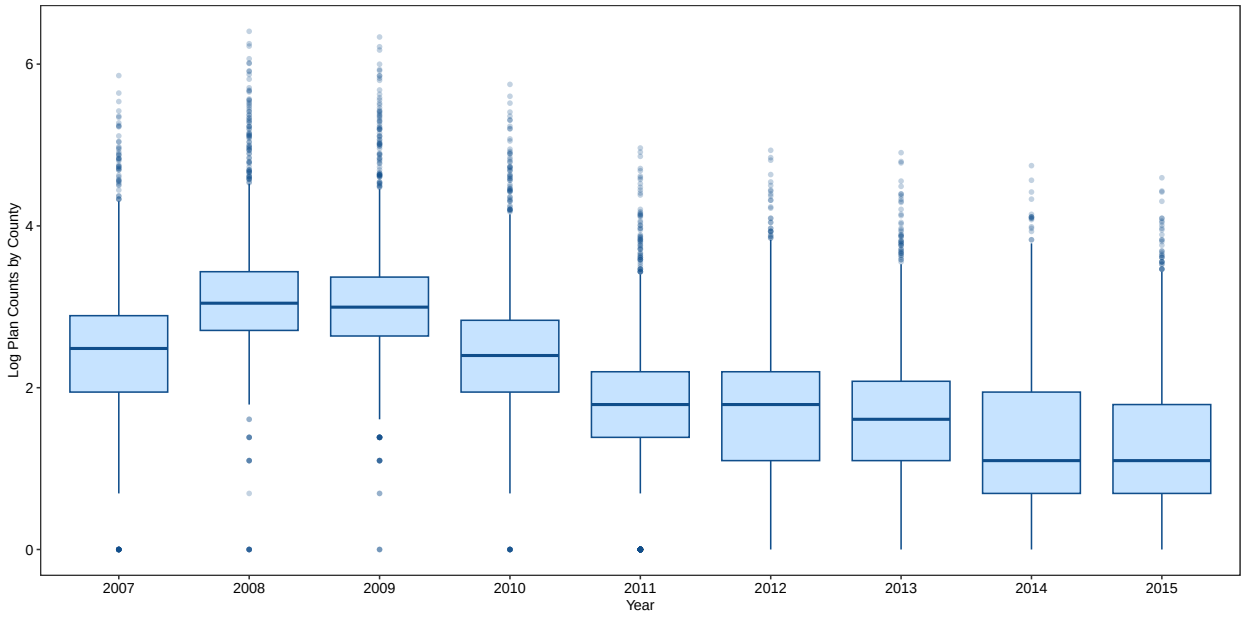


Figure 1: Distribution of Plan Counts by County from 2007 to 2015

Table 1 displays the summary statistics of plan counts by county over time from 2007 to 2015. The number of plans by county on average has been falling overtime with the sharpest decrease happening in 2011. Overall, the number of plans is too few, as 75% of counties have no more than 10 plans in any year between 2011 and 2015, while only a few counties have a relatively large number of plans.

Table 1: Summary Statistic of Plan Counts by County from 2007 to 2015

| year | Mean | Max | Q2 | Median | Q3 | Min |
|---|---|---|---|---|---|---|
| 2,007 | 17.06 | 350 | 7 | 12 | 18 | 1 |
| 2,008 | 31.82 | 605 | 15 | 21 | 31 | 1 |
| 2,009 | 29.77 | 564 | 14 | 20 | 29 | 1 |
| 2,010 | 16.47 | 314 | 7 | 11 | 17 | 1 |
| 2,011 | 8.75 | 143 | 4 | 6 | 9 | 1 |
| 2,012 | 8.05 | 139 | 3 | 6 | 9 | 1 |
| 2,013 | 7.36 | 135 | 3 | 5 | 8 | 1 |
| 2,014 | 5.79 | 115 | 2 | 3 | 7 | 1 |
| 2,015 | 5.35 | 99 | 2 | 3 | 6 | 1 |

### 2.0.2 Question 2. Provide bar graphs showing the distribution of star ratings in 2009, 2012, and 2015. How has this distribution changed over time?

Figure 2 and Figure 3 display the distribution of star ratings in 2009, 2012, and 2015. The distribution changed from right skewed to left skewed, as the majority of plans moved from the lower end of star ratings to the upper end of star ratings. The mode of the distribution shifted upwards, as most plans are rated 2.5 stars, 3 stars, and 4 stars in in 2009, 2012, and 2015 respectively. This means that the quality of the plans as measured by star ratings has improved overtime.

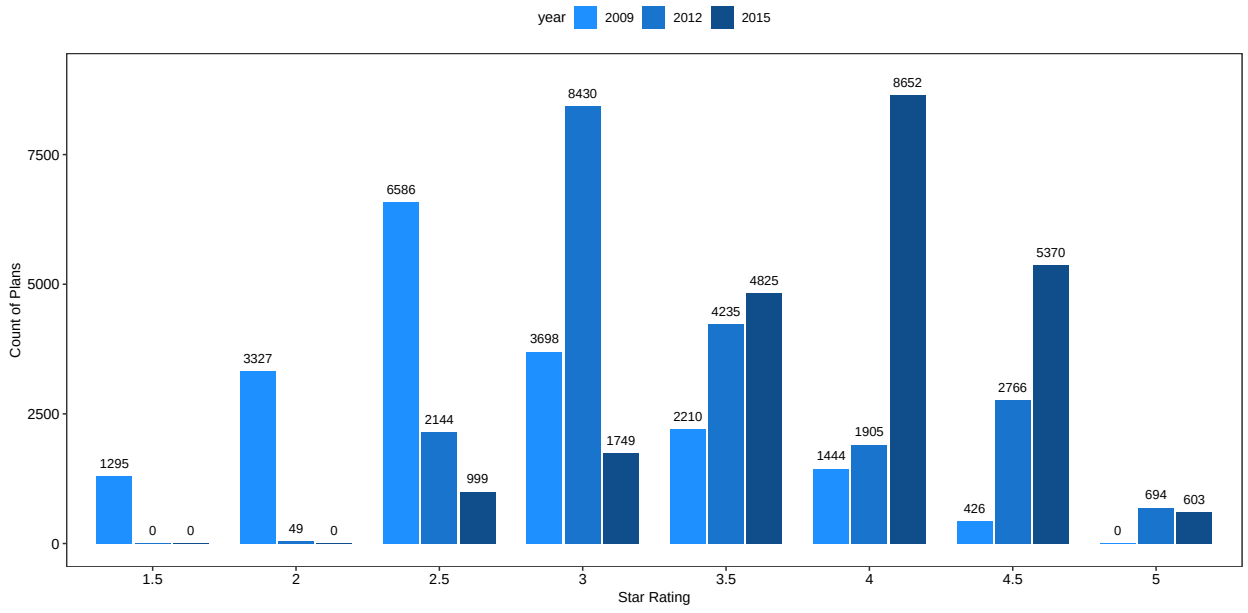


Figure 2: Frequency Distribution of Star Ratings in 2009, 2012, and 2015 (One Plot)

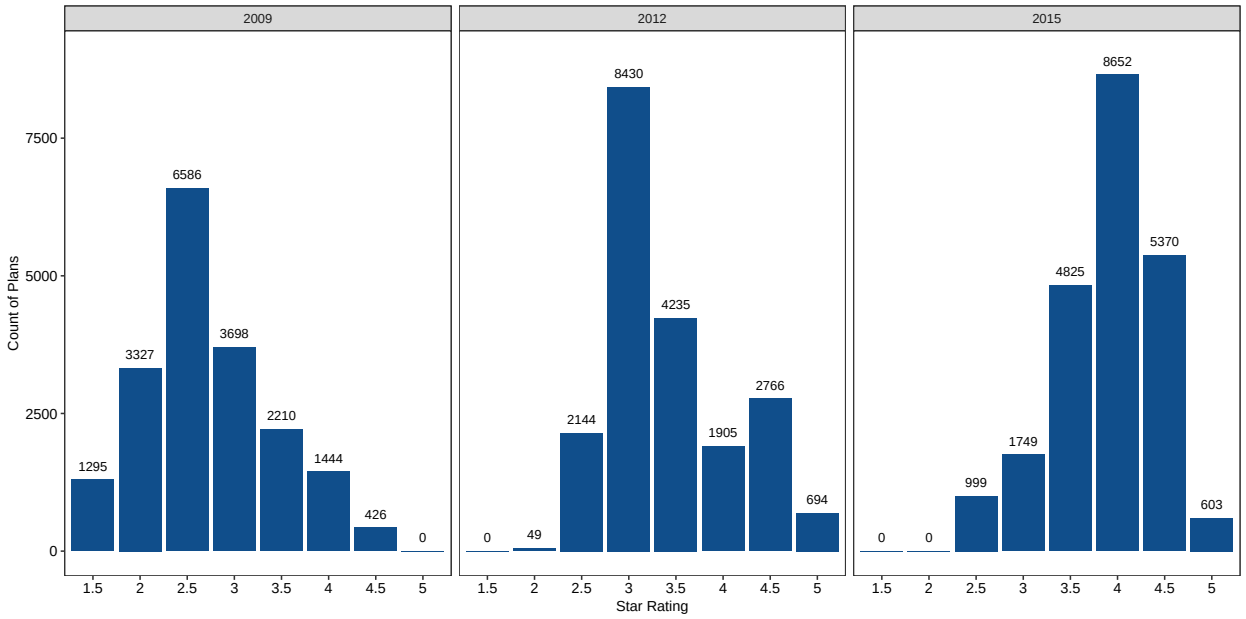


Figure 3: Frequency Distribution of Star Ratings in 2009, 2012, and 2015 (Separate Plots)

### 2.0.3 Question 3. Plot the average benchmark payment over time from 2009 through 2015. How much has the average benchmark payment risen over the years?

Figure 4 displays the average benchmark payment of Part C over time from 2009 through 2015.

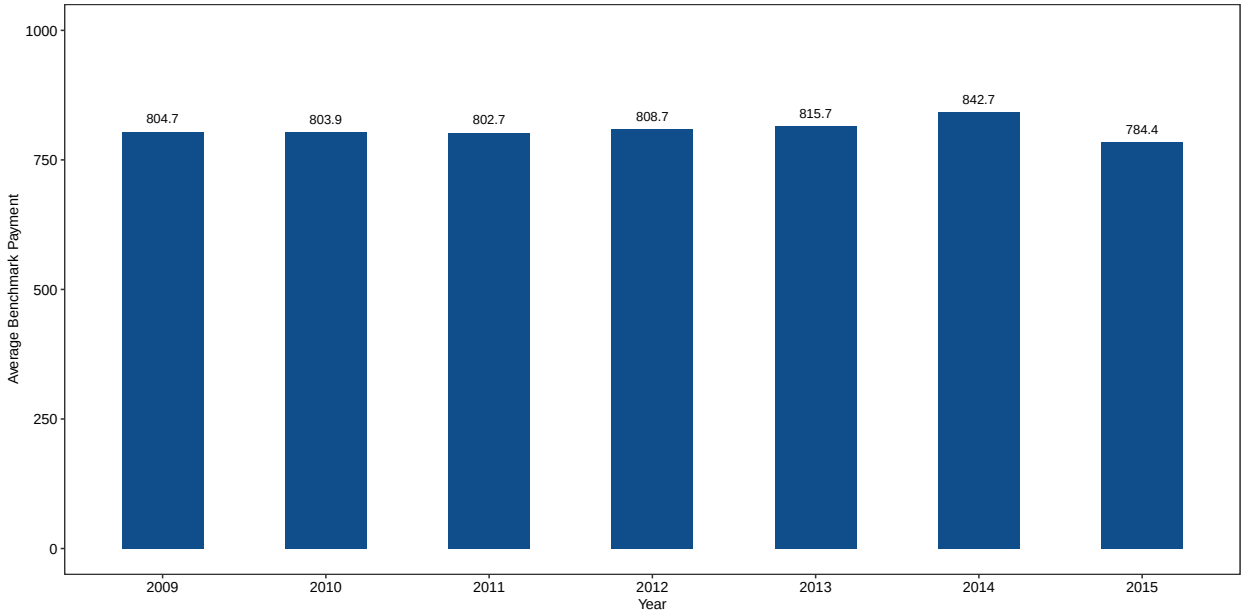


Figure 4: Average Benchmark Payment from 2009 to 2015

#### 2.0.4 Question 4. Plot the average share of Medicare Advantage (relative to all Medicare eligibles) over time from 2009 through 2015. Has Medicare Advantage increased or decreased in popularity? How does this share correlate with benchmark payments?

Figure 5 displays the average share of Medicare Advantage (relative to all Medicare eligibles) over time from 2009 through 2015. Medicare Advantage has increased or decreased in popularity as the share has been increasing over time from 2009 through 2015.

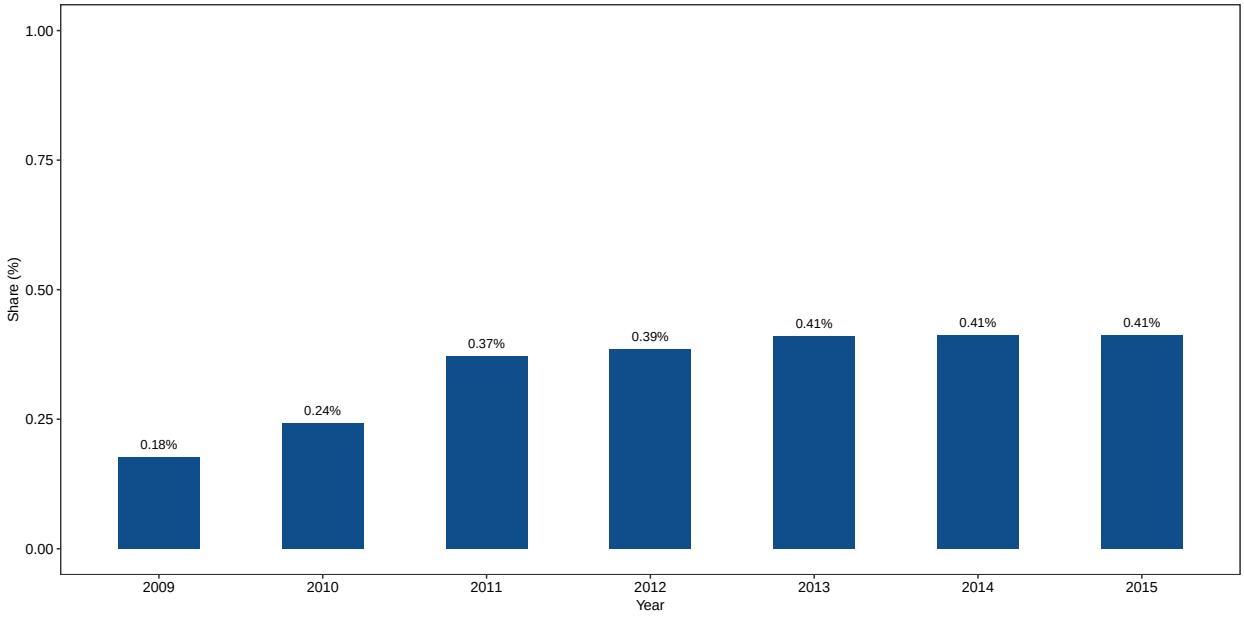


Figure 5: Average Share of Medicare Advantage (Relative to All Medicare Eligibles) from 2009 to 2015

# 3 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 2009, as this is the year in which the star rating running variable is easiest to replicate.

### 3.0.1 Question 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.

Table 2 displays the number of plans that are rounded up into a 3-star, 3.5-star, 4-star, 4.5-star, and 5-star rating.

Table 2: Number of Plans Rounded Up into Each Star Ratings in 2009

| Star Ratings | Count |
|---|---|
| 1.5 | 1,295 |
| 2 | 3,327 |
| 2.5 | 6,586 |
| 3 | 3,698 |
| 3.5 | 2,210 |
| 4 | 1,444 |
| 4.5 | 426 |

### 3.0.2 Question 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, 4 stars, and 4.5 stars. Summarize your results in a table.

Table 3: RD Estimate of the Effect of Receiving Stars at 3.5 Stars, 4 Stars, and 4.5 Stars

| | 3-Star | 3.5-Star | 4-Star |
|---|---|---|---|
| Conventional | 1,062.2258 | 1,143.1853 | 328.5270 |
| Bias-Corrected | -148.4846 | 476.4857 | -101.2723 |
| Robust | -148.4846 | 476.4857 | -101.2723 |

### 3.0.3 Question 7. Repeat your results for bandwidhts of 0.1, 0.12, 0.13, 0.14, and 0.15. Show all of the results in a graph. How sensitive are your findings to the choice of bandwidth?

Table 4: RD Estimate of the Effect of Receiving Stars at 3.5 Stars, 4 Stars, and 4.5 Stars

| | 3-Star | 3.5-Star | 4-Star |
|---|---|---|---|
| Conventional | 1,062.2258 | 1,143.1853 | 328.5270 |
| Bias-Corrected | -148.4846 | 476.4857 | -101.2723 |
| Robust | -148.4846 | 476.4857 | -101.2723 |

### 3.0.4 Question 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?

```
## $Estl
```

```
## Call: lpdensity
##
## Sample size                                      6586
## Polynomial order for point estimation    (p=)    2
## Order of derivative estimated            (v=)    1
## Polynomial order for confidence interval (q=)    3
## Kernel function                                  triangular
## Scaling factor                                   0.640377321793251
## Bandwidth method                                 user provided
##
## Use summary(...) to show estimates.
##
## $Estr
## Call: lpdensity
##
## Sample size                                      3698
## Polynomial order for point estimation    (p=)    2
## Order of derivative estimated            (v=)    1
## Polynomial order for confidence interval (q=)    3
## Kernel function                                  triangular
## Scaling factor                                   0.35952543032189
## Bandwidth method                                 user provided
##
## Use summary(...) to show estimates.
##
## $Estplot
```

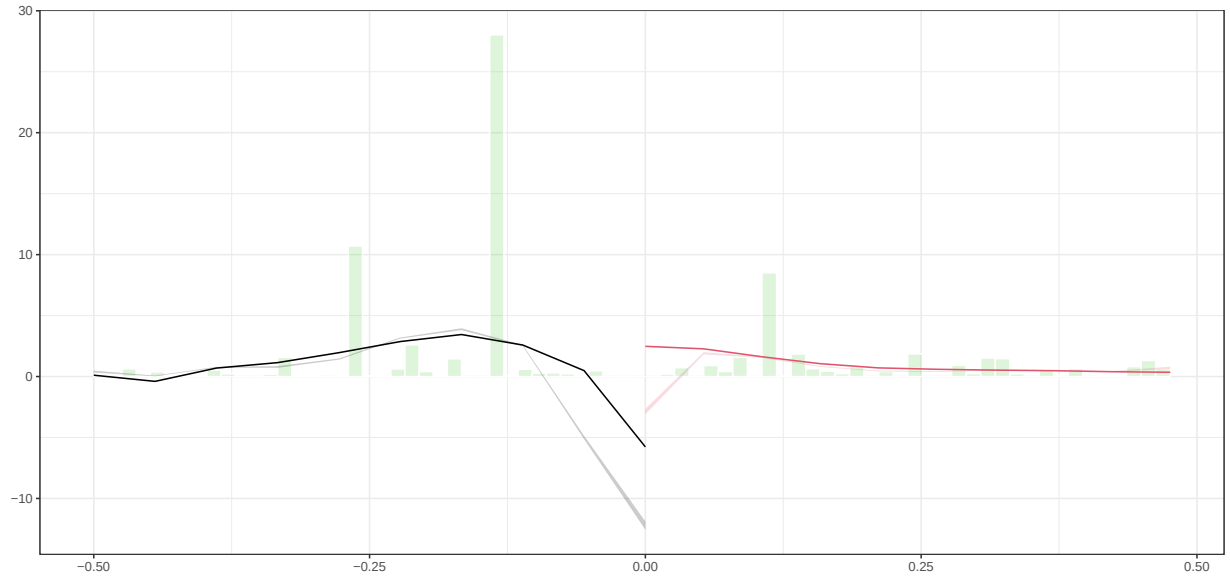


Figure 6: Distribution of the Running Variable Before and After the Relevent Threshold Values

#### 3.0.5 Question 9. Similar to question 8, 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.

```
##
## Call:
## lm(formula = avg_enrollment ~ factor(Star_Rating) + factor(partd) +
##     factor(plan_type), data = ma.data.clean)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -1327   -254   -108     41  68792
##
## Coefficients:
##                                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)                          151.761     78.865   1.924   0.0543 .
## factor(Star_Rating)2                 -88.963     49.048  -1.814   0.0697 .
## factor(Star_Rating)2.5                 9.604     50.121   0.192   0.8480
## factor(Star_Rating)3                  85.272     52.491   1.625   0.1043
## factor(Star_Rating)3.5               -61.805     61.734  -1.001   0.3168
## factor(Star_Rating)4                  68.774     67.922   1.013   0.3113
## factor(Star_Rating)4.5              -426.521     88.928  -4.796 1.63e-06 ***
## factor(partd)Yes                     178.874     27.660   6.467 1.02e-10 ***
## factor(plan_type)HMO/HMOPOS          537.885     60.704   8.861  < 2e-16 ***
## factor(plan_type)Local PPO           -79.702     67.590  -1.179   0.2383
## factor(plan_type)PFFS               -147.288     67.605  -2.179   0.0294 *
## factor(plan_type)PSO (State License)  927.288    497.000   1.866   0.0621 .
## factor(plan_type)Regional PPO        -210.115     74.138  -2.834   0.0046 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1478 on 18973 degrees of freedom
## Multiple R-squared:  0.04747,    Adjusted R-squared:  0.04686
## F-statistic: 78.79 on 12 and 18973 DF,  p-value: < 2.2e-16
```

#### 3.0.6 Question 10. Summarize your findings from 1-5. What is the effect of increasing a star rating on enrollments? Briefly explain your results.

Having a higher star rating leads to higher enrollments. Having a higher star rating signifies that the plan has a better quality which makes it more attractive to people.