Homework 4: Submission 3

Research Methods, Spring 2024

Leila Mulveny

Homework 4: Repository

Summarize The Data

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?

See Figure 1:

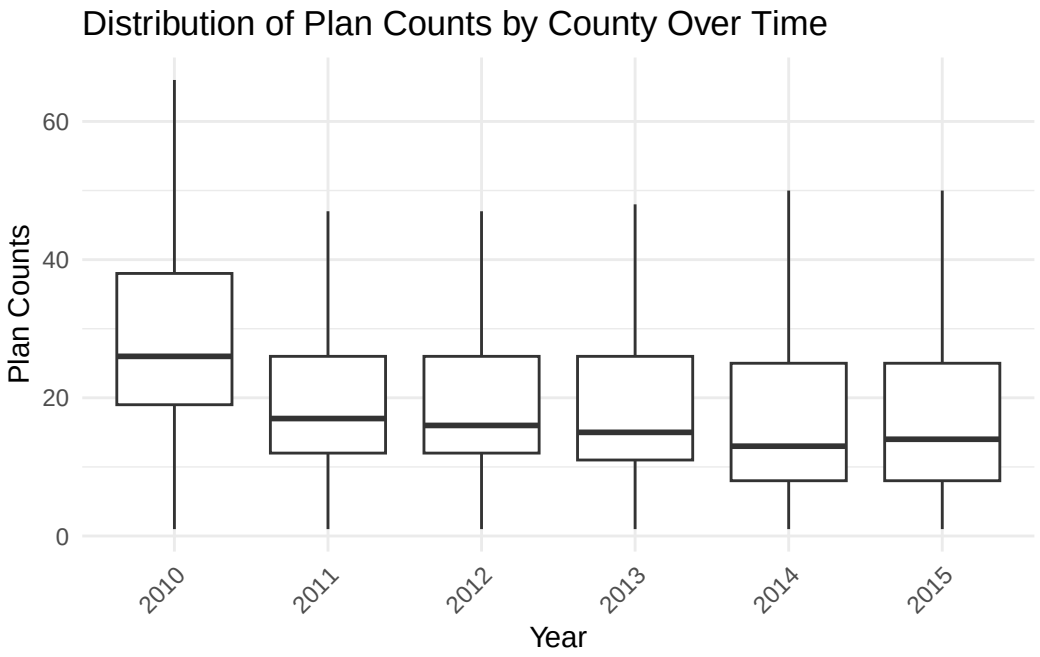


Figure 1: Distribution of Plan Counts by County Over Time

Figure 1 shows that the distribution of plan counts by country varies overtime. The median plan count is pretty close to 20 plans across the years (2010 is slightly higher). Looking at the interquartile ranges, we can see that they are relatively wide which indicates that there is a significant variation in plan counts across counties. The year 2010 stands out in that there appears to be more counties with more plans than the median—more outliers. Becuase the distribution of plan counts by county is generally consistent over time, I would say that the number of plans is sufficient.

2. Provide bar graphs showing the distribution of star ratings in 2010, 2012, and 2015. How has this distribution changed over time?

See Figure 2:

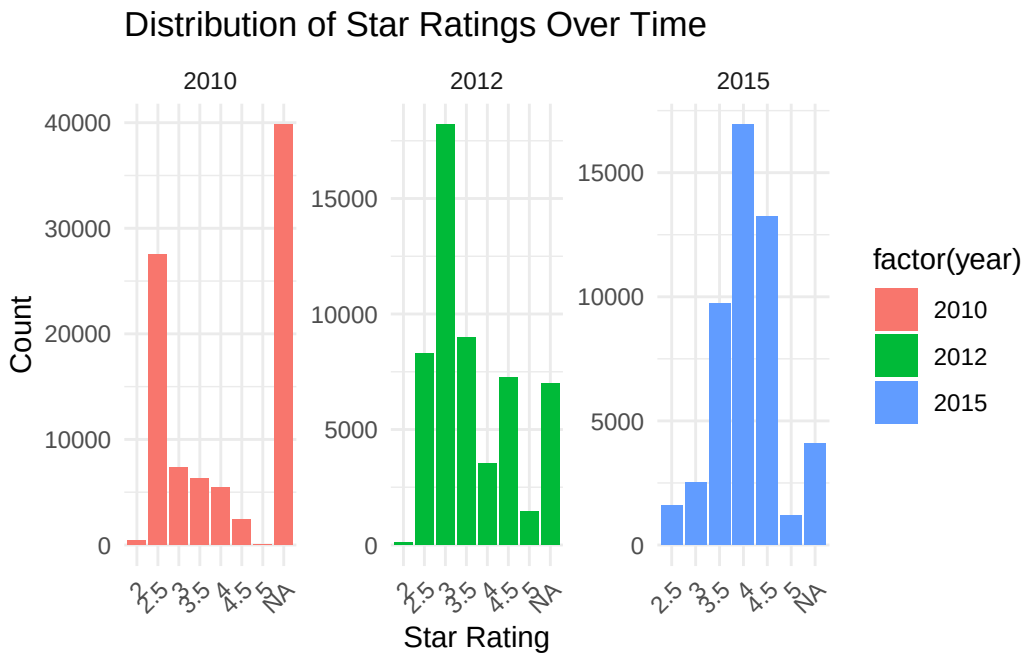


Figure 2: The Distribution of Star Ratings in 2010, 2012, and 2015

Figure 2 displays a slight shift towards higher star ratings over time. In 2010, the distribution appears more even across star ratings (2 to 4), with a clear mode around 2.5 stars. In 2015, there are more ratings of 4 and 5 stars than ever before, with the mode being around 4 stars.

3. Plot the average benchmark payment over time from 2010 through 2015. How much has the average benchmark payment risen over the years?

See Figure 3:

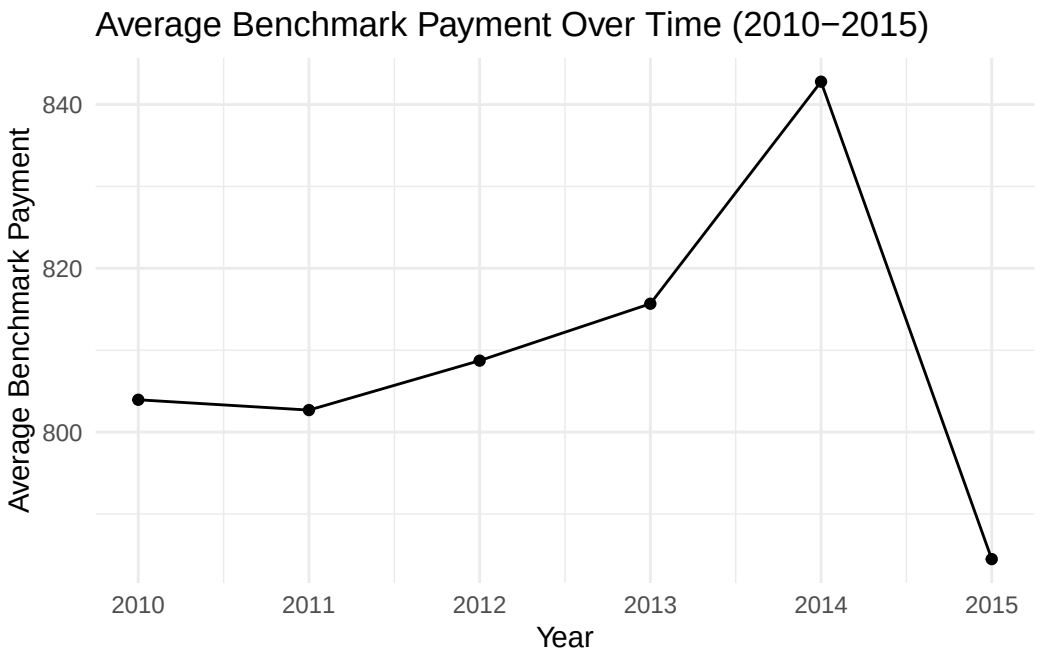


Figure 3: The Average Benchmark Payment Over Time From 2010 Through 2015

Figure 3 shows that the average benchmark payment has risen by approximately 20% between 2010 and 2015. Interestly, however, we can see a drop off in benchmark payments by about $60 in 2015 compared to 2014. This reduction is not concerning because of policy changes in 2014 are laregly attributable to this decrease. It would be interesting to look more into this decrease and any effects of policy changes in 2014 for future research.

4. Plot the average share of Medicare Advantage (relative to all Medicare eligibles) over time from 2010 through 2015. Has Medicare Advantage increased or decreased in popularity? How does this share correlate with benchmark payments?

See Figure 4:

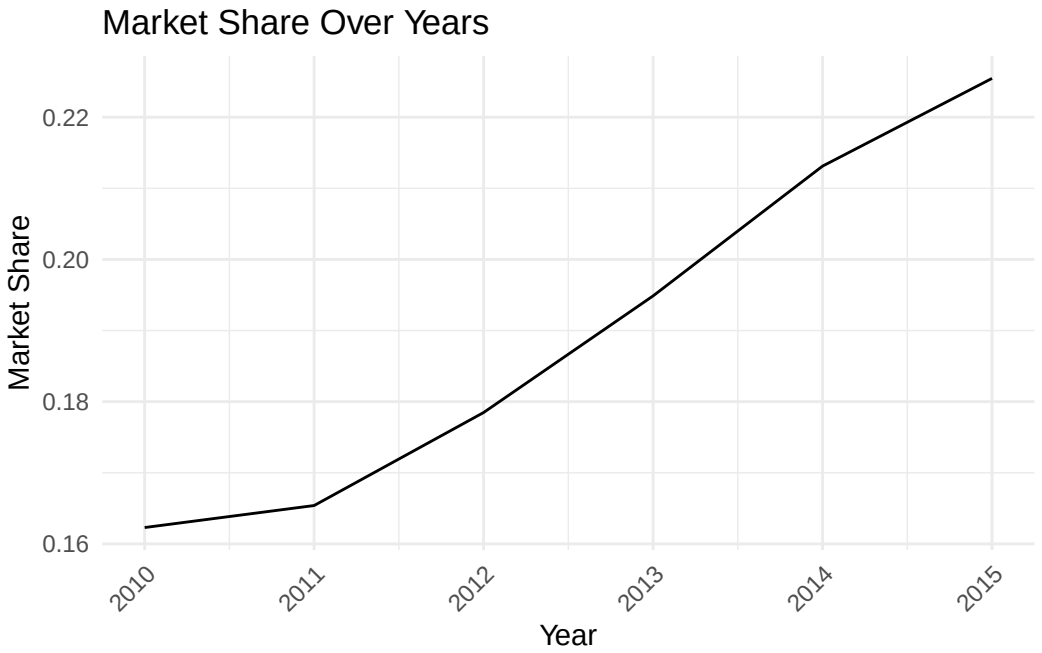


Figure 4: The Average Share of Medicare Advantage (relative to all Medicare eligibles) Over Time From 2010 Through 2015

Figure 4 shows that the average share of Medicare Advantage enrollees (relative to all Medicare eligibles) has increased in popularity over time from 2010 to 2015. This coincides with the increase in average benchmark payments over the same period (aside from the drop from 2014-2015 that was due to a policy change). Figure 4 suggests a possible correlation between the rise in benchmark payments and the growing popularity of Medicare Advantage plans. Higher benchmark payments could incentivize insurers to offer more attractive Medicare Advantage plans (benefits, coverage options, etc.), potentially leading to increased enrollment.

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.

Table 1: Number of Plans Rounded Up Into a 3-star, 3.5-star, 4-star, 4.5-star, and 5-star Rating

3-star	3.5-star	4-star	4.5-star	5-star
1734	1815	606	0	0

These results are as expected: we do not want to see any plans rounded up into a 4.5 or 5 star rating in the year 2010 as there really were not any with such a high rating. As seen in Figure 2 there are virtually no plans in 2010 with 4.5 or 5 stars and most of the distribution is at 2.5 stars. Table 1 shows us that most plans were rounded up to 3 or 3.5 stars when compared to being rounded to 4 stars or above.

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.

Table 2: RD Estimator With a Bandwidth of 0.125, Estimate of the Effect of Receiving a 3-star Versus a 2.5 Star Rating on Enrollments

	Rating	Coeff	Std..Err.	z	P.Value
Conventional	3 vs 2.5	-0.0049040	0.0033526	-1.462718	1
Bias-Corrected	3 vs 2.5	-0.0337209	0.0033526	-10.058014	1
Robust	3 vs 2.5	-0.0337209	0.0055193	-6.109645	1

The coefficient value is about -0.0049, which represents the estimated difference in enrollment probability between the 3-star and 2.5-star rating groups. A negative coefficient indicates that a 3-star rating is associated with a lower enrollment probability (intuitively not what we would expect). The effect is around 0.49%, indicating that a 3-star rating is associated with a 0.49% lower enrollment probability. However, the p-value is 1 which is greater than 0.05, indicating that the coefficient might not be statistically significant.

The downside to taking a simple average of many domains to construct the running variable is that we are finding plans with a running variable that doesn't match the actual star rating they received—as such we are getting a treatment effect that is not very accurate because the running variable is off from the true value.

Table 3: RD Estimator With a Bandwidth of 0.125, Estimate of the Effect of Receiving a 3.5-star Versus a 3 Star Rating on Enrollments

	Rating	Coeff	Std..Err.	z	P.Value
Conventional	3 vs 3.5	0.0013883	0.0029540	0.4699752	1
Bias-Corrected	3 vs 3.5	-0.0106484	0.0029540	-3.6047154	1
Robust	3 vs 3.5	-0.0106484	0.0049341	-2.1581298	1

Looking at the coefficient value for the Conventional estimate, 0.0014, we can infer that a 3.5-star rating is associated with a 0.14% higher enrollment probability compared to a 3-star rating. The coefficient here is positive but small in magnitude. However, the p-value is 1 which again suggests that the observed effect might not be statistically significant.

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?

Figure 5

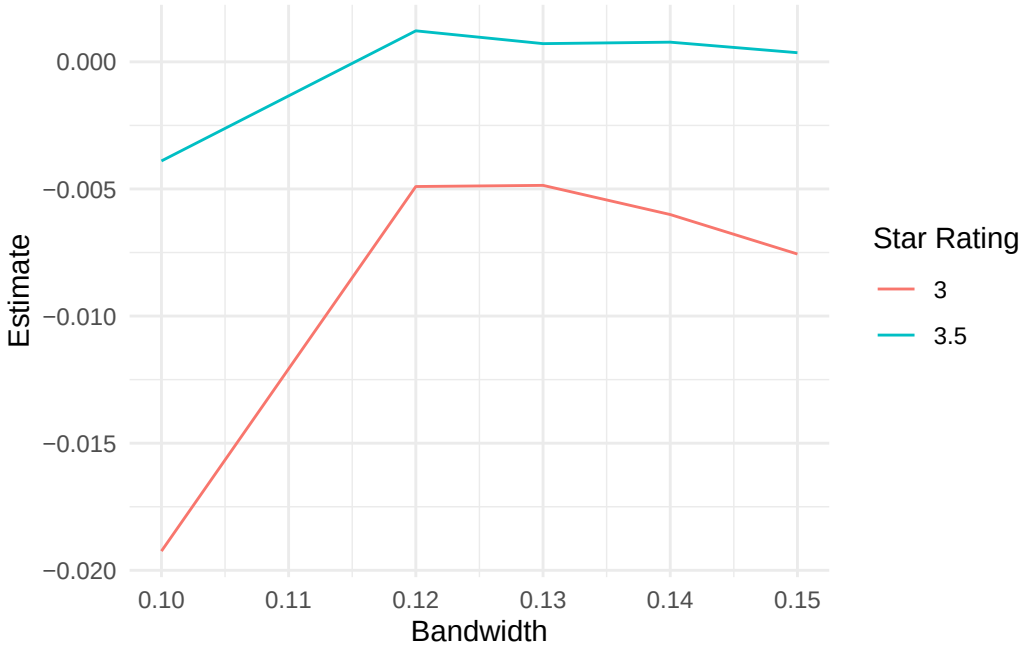
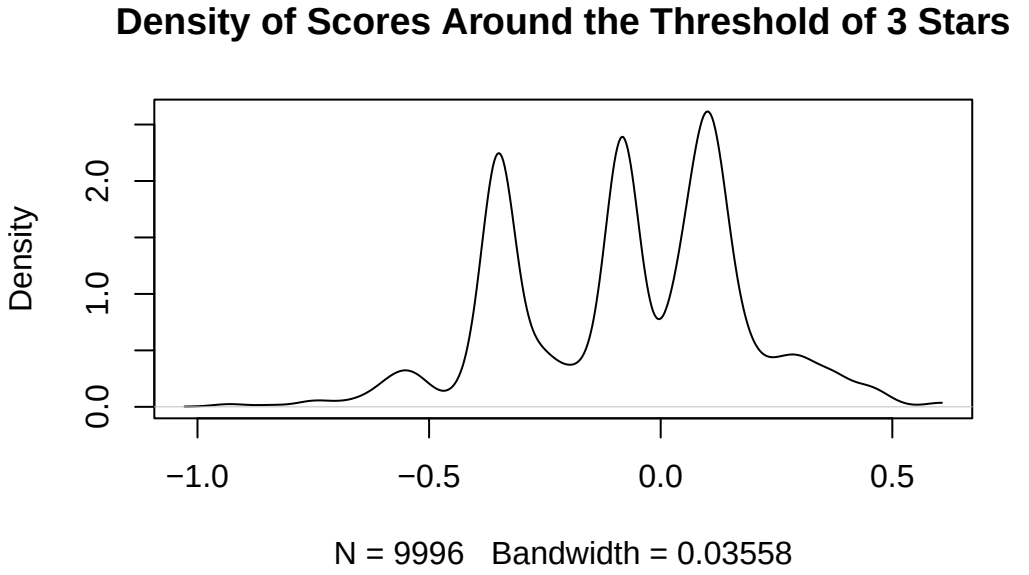
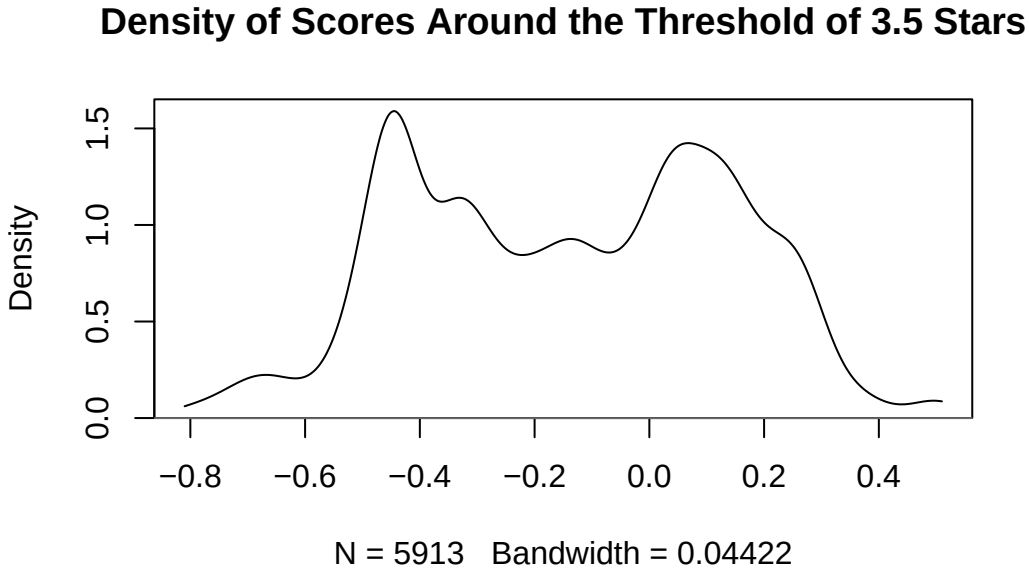


Figure 5: Sensitivity to Bandwidth: Estimated Effect of Receiving 3 Stars or 3.5 Stars Across Bandwidths

Figure 5 shows that the RD estimates are not very sensitive to the choice of bandwidth. The effect of receiving 3 stars or 3.5 stars stays relatively constant across the different bandwidths. There is a slight increase in the effect of receiving 3.5 stars compared to 3 stars as the bandwidth increases, but the overall change is small.

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?





Contracts with scores that just meet or exceed the thresholds (3 stars and 3.5 stars) are more common than contracts with scores slightly below the thresholds (more plans bunched around the 3 and 3.5 star thresholds). However, this doesn't necessarily indicate manipulation. For manipulation, we'd expect a sharp jump in scores right at the thresholds. Here, the increase is gradual, suggesting it might be a natural consequence of the rating system.

9. Similar to question 4, 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.

See Figure 6:

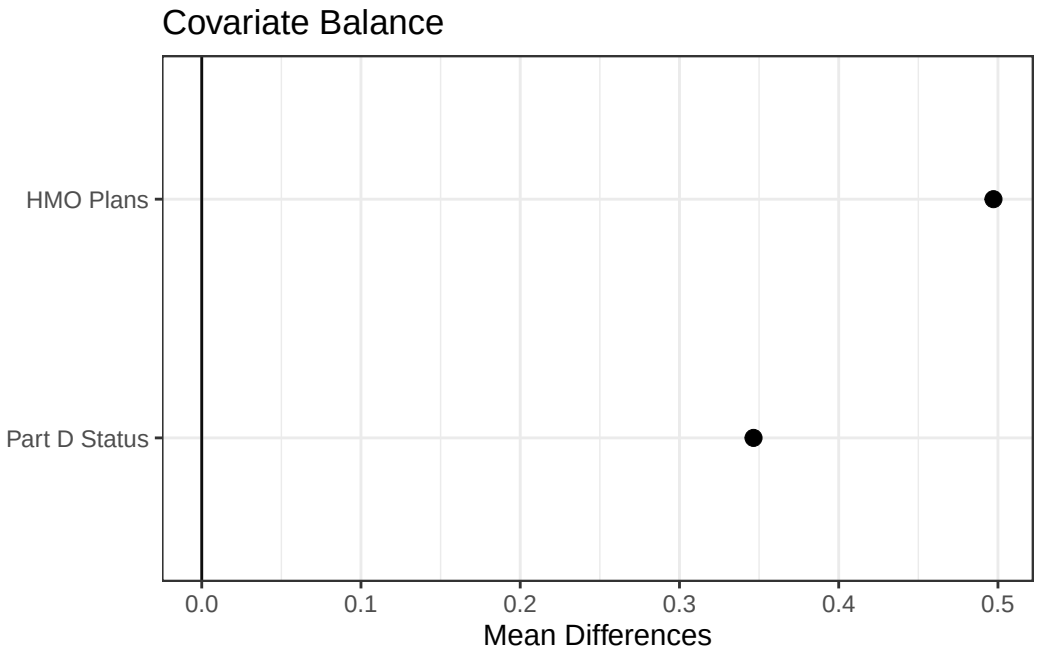


Figure 6: Comparison Below and Above 3-Star Threshold with HMO and Part D Status as Plan Characteristics

Figure 6 shows that plans with a rating above the 3 star threshold appear to have slightly different characteristics compared to plans below in terms of HMO and Part D status. The covariate balance for these characteristics is relatively small (around 0.1). So, the difference in HMO and Part D status between plans above and below the 3 star threshold is relatively small.

See Figure 7:

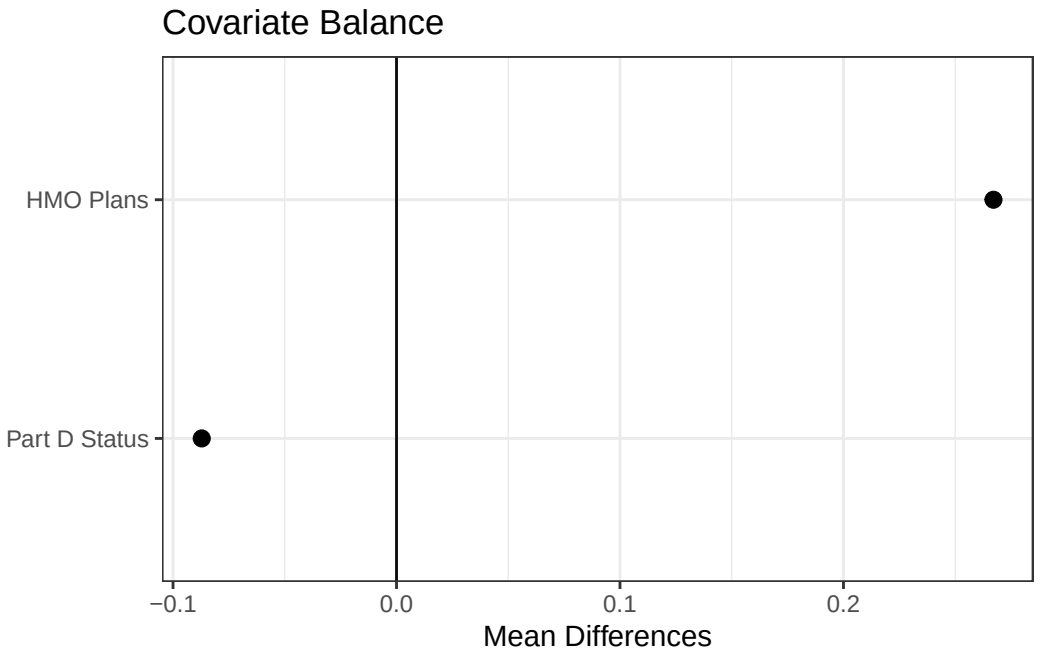


Figure 7: Comparison Below and Above 3.5-Star Threshold with HMO and Part D Status as Plan Characteristics

Figure 7 shows that plans with a rating above the 3.5 star threshold appear to have slightly different characteristics compared to plans below in terms of HMO and Part D status. The covariate balance for these characteristics is again relatively small so the difference in HMO and Part D status between plans a at the threshold is also relatively small.

10. Summarize your findings from 5-9. What is the effect of increasing a star rating on enrollments? Briefly explain your results.

The analysis suggests a positive correlation between higher star ratings and increased enrollment in Medicare Advantage plans. Questions 8 and 9 show little evidence of manipulation of star ratings, ensuring robust results. Additionally we see that people tend to avoid very low-rated plans—there is a significant drop in enrollment for plans with ratings below 3 stars. However, higher ratings don't necessarily lead to more enrollments as there is no significant increase in enrollment as the rating goes from 3 stars to higher. This means that people might not actively seek out the very best plans, but rather use the rating to avoid the worst. The biggest impact on enrollment appears to be between 2.5 and 3 stars, suggesting that 3 stars might be a critical threshold for consumers.