Name: Li Liu

Research question
What's the impact of adding verified purchase labels for reviews helpfulness?

Background
In 2015, Amazon.com introduced the "Amazon Verified Purchase" review. Customers could voluntarily add this label to their written review only if they purchased the products at Amazon.com.

Sample and population
The population is all reviews on Amazon.com. However, the reviews dataset is very large in scale and have many potential cofounders, such as product types.

The sample would be all reviews for one particular popular product across the years. There should be abundant reviews before and after the purchased labels option was introduced.

Study design
The challenge is that each review is written by different customers with various intentions. To make sure the samples in the treatment and control group are comparable, I will only a subset of matched reviews with similar length, same scores, similar sentiments, etc. Then I will use regression discontinuity to estimate treatment effect of adding the verified purchase label.

Measures
The outcome variable will be the number of helpful votes that each review received from all customers. The treatment would be adding verified purchase labels. I assume all reviewers are "compilers" as their utility would be higher with the label attached to their reviews.

Methodological challenge
Accounting for time-effect in review votes
As Amazon's reviews recommendation algorithm would favor the newer reviews with more votes, not all reviews have the same chance of being read by the customers. Also, customers wouldn't read all the reviews to make decisions. So I should make sure to adjust this problem.

Analytic method
Currently, I am planning to use the regression discontinuity. My hypothesis is that for the same (or similar) reviews, having the verified purchase label will make them look trustworthy and thus receive more votes.

External validity
I would expect my results to be applicable to the entire population of interest. However, it could be the case some customers are "never takers" as they may don't want others to know they purchased the product.