



Lab 5 - Online Reviews

Digital and Social Media Strategies

Fall 2024

Assignment Information

This formative assignment is designed to help you build and revise your knowledge of Media Mix Modelling. This material was discussed in Lecture 6. There are two components:

1. Revision Questions: Short answer questions relating to the lecture content.
2. Hands On: Use R to write short code snippets to analyse the effect of managerial responses to reviews on future ratings.

This question document, the dataset for this assignment, any additional information about the data and an R script for you to write your code are available on Canvas as a zip file "lab-05-onlinereviews.zip." Download this repository and unzip on your computer in a location where you are keeping files for this class.

We will provide solutions to the coding parts of the assignment via Canvas on Friday after the final Lab Section of the week has concluded.

You **do not need to submit this assignment for grading.**

Learning Goals

By the end of this assignment, you will be able to:

- Explain research findings in the academic literature on online reviews as they relate to impacts on product demand, the effects of review incentivization and the prevalence and effect of fake reviews
- Design a field experiment to test whether disclosing a review is incentivized impacts consumer demand for products
- Analyse data from natural experiment that investigates the effectiveness of managerial responses to online reviews on future review ratings and draw managerially relevant conclusions

Revision Questions

1. List three factors that contributed to the rapid growth in social media. Explain the contribution of each in one sentence.
2. Explain why user generated star ratings could have a causal impact on demand for experience goods. (max. 3 sentences)
3. List two alternative mechanisms that could serve the same purpose as star ratings. Explain why they don't offer the same type of information as star ratings (max. 3 sentences)
4. Explain intuitively how Luca uses discontinuities in Yelp's star ratings scheme to estimate causal effects of star ratings on demand. Interpret the estimated effect sizes and outline how they differ across restaurant types. (max. 6 sentences)
5. Why do managers decide to incentivize consumers to write reviews? (max 2 sentences)
6. What effects do Frandkin and Holtz find when Airbnb introduces review incentivization? Why is this the case? Discuss the implications for marketers and managers. (max 6 sentences)
7. Assume you work for a large online retail platform such as Zalando. Construct a field experiment that examines whether disclosing that a review is incentivised impacts consumer demand.
HINT: In your experiment design be careful about how to construct treatment and control groups so that there are no spillover effects from the treatment onto the control group.
8. How does Mayzlin and Dover and Chevalier investigate when review manipulation occurs? What are their main findings? (max 6 sentences)
9. What do fake reviews do to sales, prices and sales rank of products that pay for them? (max. 3 sentences)
10. How does He, Hollenbeck and Proserpio assess whether fake reviews are bad for consumer welfare? What do they find? (max. 3 sentences)

Hands On: Managerial Responses to Online Reviews

This exercise studies online reputation management through the use of public comments by firms in response to online reviews. The content is based on the article "Online reputation management: Estimating the impact of management responses on consumer reviews" by Proserpio and Zervas. The article is published in Marketing Science in 2017, and is available in our course readings. The paper wants to investigate the relationship between a hotel's use of management responses and its online reputation (measured by star rating) & establish a

causal relationship from the use of management responses to online reputation. Your goal in this exercise is to explain key arguments and replicate selected results from this paper. The data for this exercise is located “data/responses.dta”. We recommend you read the paper carefully as you work through the exercises.

Task 1: Setting Up Your R Session

- Open R.
- Start a new RStudio project in the folder “lab-05-onlinereviews” that you have already downloaded unzipped (see page 1).
- Open the script file “managerial_reponses.R”.

Task 2: Explaining the Empirical Approach

- Explain why firms might use public responses to reviews to manage their online reputation. Should they respond relatively more to positive or negative reviews? Explain why.
- Proserpio & Zervas' empirical exercise uses what they call 'cross-platform' difference in differences. Using your own words, explain their idea conceptually - and justify why it is valid. You can use equations or figures, but do so sparingly. (max 7 sentences).
- Explain what the 'parallel trends' assumption is. Why is it important in this application? Which figure (if any) provides support for the parallel trends assumption?
- What is the 'Ashenfelter dip'? Why do the author's believe they see a pattern akin to an Ashenfelter dip in their application.

Task 3: Doing Difference in Differences: Computing Differences in Means

- Load the data for this exercise and name it “hotels_orm”. For this exercise you will only need the rows where `xplatform_dd_obs = 1`. Keep only the columns hotel_id, year, stars, after, ta_dummy, first_response, cum_avg_stars_lag, log_count_reviews_lag, t, and ash_interval.
- Create a data frame with two rows and two columns where the rows take the values of `first_response` = 0 or 1, and the columns take the values of `ta_dummy` = 0 or 1. The values in the data frame should be the respective group means of `stars`.

HINTs:

- first_response is the variable that tells us whether a review is before (first_reponse = 0) or after (first_response = 1) the first managerial response by a hotel.
- ta_dummy tells us whether the review is posted on Trip Advisor (ta_dummy =1) or Expedia (ta_dummy =0)

This constructs the familiar Difference in Differences table:

	Before (first_reponse =0)	After (first_reponse =1)
Control Group (ta_dummy = 0)	β_0	$\beta_0 + \beta_1$
Treatment Group (ta_dummy = 1)	$\beta_0 + \beta_2$	$\beta_0 + \beta_1 + \beta_2 + \delta$

- C. Compute the difference between `first_response` = 1 and `first_response` = 0 for each of `ta_dummy` = 0 and `ta_dummy` = 1.
- D. Compute the difference between the two values in C to get your DiD estimate.
- E. Interpret the estimate that you obtain

Task 4: Doing Difference in Differences: Regression Implementation

Use the following regression equation to estimate the difference in difference estimator of the effect of managerial responses on online reputation:

$$Stars_{ijt} = \beta_1 After_{ijt} + \beta_2 TripAdvisor_{ij} + \delta After_{ijt} \times TripAdvisor_{ij} + \varepsilon_{ijt}$$

where $Stars_{ijt}$ is the star-rating of review i for hotel j in calendar month t , $After_{ijt}$ is an indicator for reviews (on either platform) submitted after hotel j started responding, $TrippAdvisor_{ij}$ is an indicator for TripAdvisor ratings and ε_{ijt} is the error term. The relationship between the variables in the equation above and the variables in the dataset is:

- After is the variable `first_reponse` in the data,
- TripAdvisor is the variable `ta_dummy` in the data, and
- After X TripAdvisor is the variable called `after`

Let's estimate this model:

- A. Estimate the "simple" Difference in Differences regression model and report the results
- B. Re-estimate the model in A, but omit observations where `ash_interval` = 0 to correct your estimates for the Ashenfelter Dip described in Task 2. Report the results.

Task 5: Interpreting the Results & Drawing Conclusions

- A. Interpret the value of the coefficients you deem most important from Task 4. Are these effects significant from a marketing perspective (i.e. Should they shape marketing practice)?
- B. Explain intuitively why management responses to reviews can lead to improved hotel ratings. Can you support this argument using any of the other results in Proserpio and Zervas' work?
- C. Based on your results, should managers respond to online reviews?
- D. Are there other outcome variables besides future star ratings you would like to see the results of to make a more definitive conclusion? Explain your answer.