Assessing Online Platform's Policy Impact on Review Manipulation

COMM635 Advanced Topics in Information Systems

UBC Sauder School of Business

Instructor: Arslan Aziz
Student Name: Eugenie Lai
Student Number: 38920154

May 26, 2020

1 Introduction

Preferred by more than half of the population (Statista, 2018), online retailing is one of the fastest-growing sectors in the past decade. 89% of consumers browse online reviews prior to making buying decisions (Hu et al., 2011). However, studies have shown that unnatural reviews take up nearly a third of total online reviews. Moreover, review manipulation is shown to have substantial impacts on consumer's buying decisions (Lappas et al., 2016; Che and Hörner, 2017). Therefore, there is a growing need to monitor, assess, and control the quantity and impact of unnatural reviews on online platforms.

The three main types of unnatural reviews are negative reviews by competitors, positive reviews by producers, and incentivized reviews by customers. As a method of review manipulation, incentivized reviews are posted by consumers who received incentives from the product seller for the purchased product. Some common incentives include free products, product discounts, promotion coupons, and cash backs.

Amazon is the largest online retailer subscribed by 82% of American households by June 2019 (Consumer Intelligence Research Partners, 2019). Amazon supports an online retailing platform for millions of sellers across 29 sectors, ranging from home and kitchen to sports and outdoors. Besides providing a platform for businesses, Amazon also has been featuring its own brands across over ten sectors including clothing, homeware, electronics, and office products.

Historically, Amazon has prohibited compensation for reviews but allowed businesses to incentivize their customers to share their "honest" opinion as long as the reviewers disclosed their affiliation with the business in their review text. Although, in theory, these reviewers could write their unbiased opinion on the product, studies have shown that these incentivized reviews have tended to be overwhelmingly biased in favour of the product being rated. A 2016 study by ReviewMeta analyzed over seven million reviews on Amazon and indicated that the average rating for products with incentivized reviews was 0.38 stars higher than non-incentivized ones. The 0.38-star difference is able to boost products from the 54^{th} percentile to the 94^{th} percentile. In addition, the study also found that incentivized reviewers were 12 times less likely to give a 1-star rating than non-incentivized reviews. Hence effectively, incentivized reviews could create top-rated products.

In an effort to keep bias out of the review process, on October 6, 2016, Amazon announced to ban all incentivized reviews across the platform. In this study, we seek to examine whether review policy bans can reduce review manipulation on online platforms. Accordingly, we seek to answer the following research question:

> *Did the ban on incentivized reviews have an impact on review manipulation? Specifically, did the ban change the nature and number of incentivized reviews?*

To answer the research question and guide our empirical examination, we integrated the literature on the impact of online review manipulation in information systems and unnatural review detection in computer science. The primary contribution of this study is to provide a proof-of-concept for capturing the impact of the review policy ban on incentivized reviews across Amazon.

2 Related Literature

2.1 Study Context

Amazon announced to eliminate any incentivized reviews on October 6, 2016. This provides us with a natural shock as an exogenous variation that only directly affects incentivized reviews.

In our data collection process, we affiliated with ReviewMeta, an independent, non-profit site that helps online shoppers obtain unbiased product ratings on a number of online retailing platforms including Amazon and Bodybuilding.com. The ReviewMeta site allows shoppers to input a product URL, and it will examine reviews of the given product and then create a report by running an analysis of the reviews.

2.2 The Impact of Online Review Manipulation

Online reviews can be a useful information source (Lu and Rui, 2018). In the field of information systems, the emerging literature on online reviews and platform policy has primarily focused on assessing the impact of online reviews and review manipulation. Previous studies have shown that review platform penetration leads to economically quality improvements (Zhao et al., 2015), and the proliferation of online reviews can increase competition in markets where quality is not fully revealed at the time of purchase (Zhao et al., 2015).

More importantly, many previous studies have established the substantial impact of online reviews on consumers' buying decisions (Chevalier and Mayzlin, 2006; Mayzlin et al., 2014; Luca, 2011; Lappas et al., 2016). 89% of consumers would consider online reviews (Goh et al., 2013), and Most people rely more on average review scores and are willing to settle for items with lower average scores if items are more popular (Analytis et al., 2017).

However, unnatural reviews comprise up to a third of the total online reviews and are more likely to appear in competitive markets (Jindal and Liu, 2008; Hu et al., 2011), and some forms of review manipulation are more powerful than others (Lappas et al., 2016; Che and Hörner, 2017; Zhuang et al., 2018; Ivanova and Scholz, 2017). There are three types of unnatural reviews: negative reviews by competitors, positive reviews by producers, and incentivized reviews by customers. In practice, unnatural reviews are used to manipulate consumers' buying decisions and could cause damages to the platform. For example, unnatural ratings are used to increase sales of low quality, and self-injecting positive reviews are used to increase product popularity. The degree of influence depends on the consumer as well. Consumers with less online shopping experiences may fall victim to unnatural reviews. However, the effect of unnatural ratings is not pronounced if retailers are selling high-quality products.

Individuals' subjectivity and biases are also influencing online reviews besides product quality and review manipulation. Social influence leads to disproportionately positive online ratings, and subsequent raters more likely to be influenced by previous positive ratings than negative ones (Aral, 2014). In addition, online product ratings are a better tell of customer satisfaction rather than quality, and the baseline of online ratings is built by expectations (Engler et al., 2015). Gadidov and Priestley (2018) shows that Yelp reviews relevant to operational performance and evaluation of customer satisfaction.

2.3 Unnatural Review Detection

A rich stream of literature (Barbado et al., 2019; Heydari et al., 2016; Rayana and Akoglu, 2015; Wang et al., 2016, 2017) in computer science has been exploring the topic of unnatural review detection since the 2000s. Examples of the popular methods are deep learning models, graph-based models, and probability models (e.g., Markov model). Features used in unnatural review detection have also been extensively explored. Commonly used features are linguistic features of the reviews and behavioural and network features of the reviewers. Many studies (Rayana and Akoglu, 2015; Wang et al., 2016, 2017) found that those methods outperform human experts, and unnatural reviews take up roughly 15% of total reviews on online platforms. However, despite the extensive work in the area, the accuracy of the state-of-the-art approaches is less than 90%. Hence with the existing tools, it is still challenging to identify whether a review is unnatural or not without knowing the ground truth.

3 Data

ReviewMeta is an independent site that helps consumers get a better understanding of Amazon reviews. By affiliating with ReviewMeta, we collected a small dataset of reviews on 2 categories of top-selling products: tablets and charging cables. The dataset was collected based on lists of top-selling products in charging cables and tablets in September 2019. The dataset consists of reviews for Amazon and non-Amazon products six months before and after the policy ban. It contains all the raw variables present in the Julian dataset and additional variables created by ReviewMeta. For example, stated incentivized reviews are detected by ReviewMeta's own methodology, assuming the methodology stays consistent within the one-year time frame.

Variable	Obs	Mean	Std.Dev.	Min	Max
Incentivized	93287	.002	.047	0	1
Verified	93287	.959	.198	0	1
Positive	93287	.641	.407	0	1
Rating	93287	4.209	1.279	1	5
Word Count	93287	30.903	52.279	0	2676
Helpfulness	93287	.006	.147	0	21
Image Count	93287	.002	.081	0	7
After Ban	93287	.539	.498	0	1
Weekly Reviews	93287	1569.539	1540.867	1	6007

Table 1: Descriptive Statistics of the ReviewMeta dataset

We analyzed the review sentiment using Amazon Web Service (AWS) Comprehend, which gives a sentiment score for positive, negative, neutral, and mixed sentiment for every review. We added the positive sentiment score to the dataset and used along with the rating as another measurement of reviews. We use this dataset to obtain a proof-of-concept by leveraging the labels of the stated incentivized reviews.

Our collected data has three tablet brands and 18 phone cable brands. The two product categories have a total number of distinct products of 17 and 24 respectively. We intended to collect review data for the 100 top-sellers in each category. Due to the variation in product characteristics (e.g., colour, size), a product can have a number of product IDs sharing the same review section. Hence we obtained 17 distinct tablets and 24 distinct phone cables with the list of 100 top-sellers. The number of brands also reflects the competitiveness of the market (Zhuang et al., 2018). Although the number of top-sellers is similar across the two categories, we have six times more brands in phone cable. This indicates an important time-invariant characteristic of the product categories for which we need to impose additional control.

4 Analysis

4.1 Research Questions

The causal question we seek to answer and our hypothesis are detailed below.

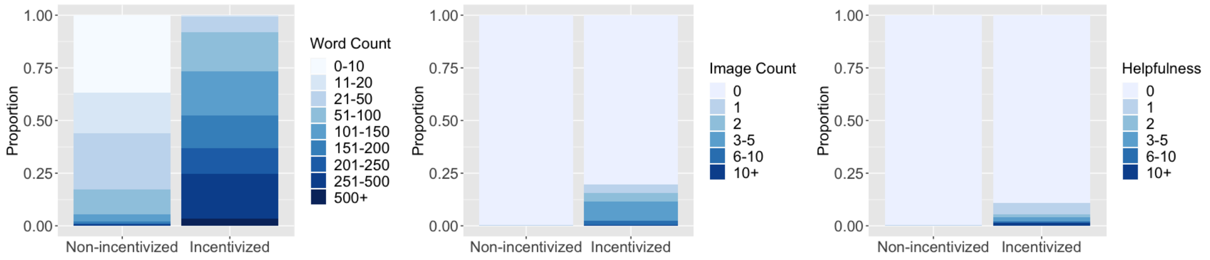
> *Did the ban on incentivized reviews have an impact on review manipulation? Specifically, did the ban change the nature and number of incentivized reviews?*
>
> **Hypothesis:** After the ban, the nature of reviews changed, and the number of incentivized reviews increased across the platform. The stated incentivized reviews would go down. They would either get deleted by the platform or become unstated incentivized reviews. As a result, unstated incentivized reviews will increase, which makes them challenging to detect. Hence we would observe an increase in unidentifiable incentivized reviews after the ban, and therefore, review manipulation would become more severe

In summary, we argue that after the policy ban, there are more incentivized reviews, and they become more similar to natural reviews.

4.2 Methodology

4.2.1 Empirical Model

As Section 3 describes, we collect data from ReviewMeta. The ReviewMeta dataset has two main characteristics: (1) it only includes reviews for the top sellers in the cable and tablet category; (2) ReviewMeta flagged the stated incentivized reviews with their NLP methods. We choose the ReviewMeta



(a) (b) (c)

Figure 1: Proportion of three review characteristics of incentivized and non-incentivized reviews: (a) Incentivized reviews are longer; (b) Incentivized reviews have more images attached to them; (c) Incentivized reviews appear to receive more helpful yes's than non-incentivized reviews.

dataset to produce a proof-of-concept in this study to take advantage of the stated incentivized review labels.

Here we take a difference-in-difference approach: for each product category, we examine the difference in the nature of reviews across Amazon and non-Amazon products and across time periods before and after the policy ban. In the first difference, we use Amazon products as the control group as Amazon itself would not incentivize its customers to post reviews. In this study, we will use a one-year time frame, including six months before and after the ban.

In addition to the average policy effects on the platform, our second question is interested in the potential heterogeneous effects in different product categories across the platform. Since different product categories have their unique time-invariant characteristics, we use fixed effects on the category level and brand level in addition to the difference-in-difference analysis described above.

Furthermore, we need to check if the assumptions of using the difference-in-difference approach are satisfied. Here we check the pre-treatment parallel trend assumption as the following. The interaction term of the weeks before the ban are tested. We create a dummy (i.e., 26 levels) for each week and then take the interaction of the non-Amazon dummy with this variable. The pre-treatment parallel assumption is satisfied if all the coefficients are statistically zero.

Lastly, our future plan for additional robustness checks will be in Section 6.

4.2.2 Key Measures

The ground truth of whether a review is incentivized or not is almost impossible to obtain. We need other dependent variables to observe the changes in incentivized reviews. Hence we explore other observable characteristics to capture the policy effects on incentivized reviews

Although it is impossible to identify the ground truth of unstated incentivized reviews, stated incentivized reviews can be identified using NLP methods. In our analysis, we assume that the stated incentivized reviews prior to the ban share the same characteristics as the unstated incentivized reviews both before and after the ban. We use the stated incentivized reviews identified in the ReviewMeta dataset and did data exploratory analysis on review length, image count, and helpfulness. We find that incentivized reviews are longer, with a higher image count, and more helpful than non-incentivized reviews (Figure 1). Additionally, there is a strong correlation between reviews' helpfulness and length and image count. We validate what we observe in the data using regression tests with fixed effects on product categories and brands (Figure 2) and find that those characteristics are statistically significant for incentivized reviews. Along with review rating and sentiment, these five observables are the dependent variables in the difference-in-difference analysis

	Dependent variable:				
	Rating	Sentiment	Length	Helpfulness	Image count
	(1)	(2)	(3)	(4)	(5)
Incentivized	0.325***	0.122***	150.925***	0.156***	0.244***
	(0.088)	(0.028)	(3.598)	(0.010)	(0.006)
Observations	93,287	93,287	93,287	93,287	93,287
R^2	0.0001	0.0002	0.019	0.003	0.020
Adjusted R^2	0.0001	0.0002	0.018	0.002	0.020
F Statistic (df = 1; 93284)	13.632***	18.995***	1,759.606***	234.074***	1,932.406***
Note:					*p<0.1;* ***p<0.05;*** p<0.01

(a)

	Dependent variable:				
	Rating	Sentiment	Length	Helpfulness	Image count
	(1)	(2)	(3)	(4)	(5)
Incentivized	0.203**	0.074**	149.417***	0.167***	0.253***
	(0.090)	(0.029)	(3.677)	(0.010)	(0.006)
Observations	93,287	93,287	93,287	93,287	93,287
R^2	0.0001	0.0001	0.017	0.003	0.021
Adjusted R^2	-0.0002	-0.0002	0.017	0.002	0.021
F Statistic (df = 1; 93265)	5.087**	6.626**	1,651.351***	254.245***	1,978.199***
Note:					*p<0.1;* ***p<0.05;*** p<0.01

(b)

Table 2: Regression with (a) Category-Level Fixed Effects; (b) Brand-Level Fixed Effects

5 Results

5.1 Main Effects

The results align with our hypotheses. Based on our difference-in-difference analysis without fixed effects (Figure 3), there is a statistically significant decrease in review length, image count, and helpfulness for non-Amazon products after the ban. This means that after the ban, reviews became shorter, have fewer images, and are less helpful. Hence the policy ban had an impact on the nature of reviews. Notably, the coefficient of review rating and sentiment is positive but not significant

In our difference-in-difference analysis with category-level fixed effects (Table 4), besides the treatment effects being consistent for all dependent variables, review rating and sentiment also become statistically significant. The impact of the policy ban on incentivized reviews becomes more pronounced since we control the time-invariant characteristics of the product categories. In our previous analysis, we find that incentivized reviews inherently have higher ratings and sentiment. Therefore, the statistically significant increase in review rating and sentiment after the ban signals an increase in incentivized reviews. And instead of reducing the number of incentivized reviews on the platform, the policy ban led to more natural-looking incentivized reviews.

Lastly, the result of our difference-in-difference analysis with brand-level fixed effects (Table 5) is consistent with our previous findings. It is expected that we lose the statistical significance of the dependent variables as we further break down the data.

Our key findings are summarized as the following. First, the review policy ban changed the nature of reviews. Reviews become shorter with fewer images attached. Second, the policy ban increased review rating and sentiment across product categories. Third, our evidence shows that the ban may have heterogeneous effects on different product categories due to their individual characteristics.

	Dependent variable:				
	Rating	Sentiment	Length	Helpfulness	Image count
	(1)	(2)	(3)	(4)	(5)
Non-Amazon product	0.143***	0.043***	0.045	0.015***	0.011***
	(0.041)	(0.012)	(1.439)	(0.005)	(0.004)
After ban	-0.076	-0.012	-4.688***	-0.010*	-0.002
	(0.050)	(0.015)	(1.785)	(0.006)	(0.005)
Non-Amazon product x After ban	0.071	0.021	-4.000**	-0.015**	-0.011**
	(0.057)	(0.016)	(2.011)	(0.006)	(0.005)
Constant	4.131***	0.618***	34.661***	0.010**	0.002
	(0.035)	(0.010)	(1.257)	(0.004)	(0.003)
Product Category Fixed Effects	No	No	No	No	No
Brand Fixed Effects	No	No	No	No	No
Observations	3,146	3,146	3,146	3,146	3,146
R^2	0.013	0.014	0.031	0.022	0.010
Adjusted R^2	0.012	0.013	0.030	0.021	0.009
Residual Std. Error (df = 3142)	0.645	0.187	22.906	0.073	0.058
F Statistic (df = 3; 3142)	14.035***	15.048***	33.733***	23.992***	10.944***
Note:					*p*<0.1; ***p*<0.05;** p<0.01

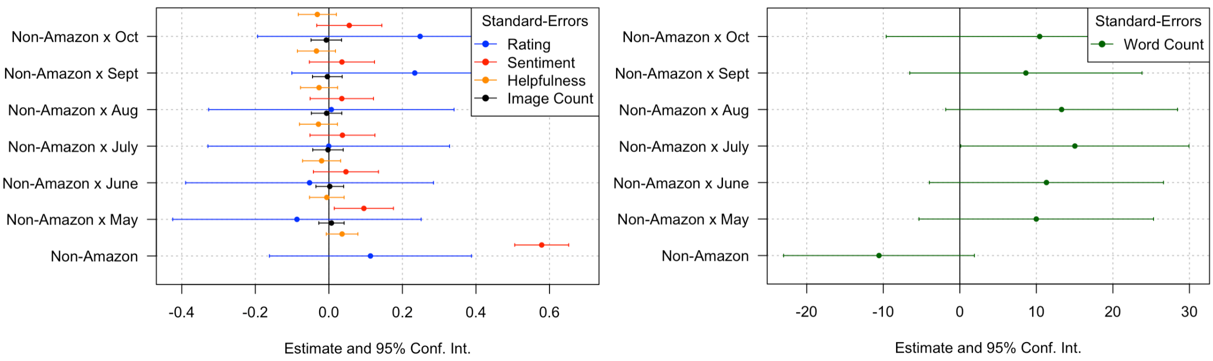
Table 3: Difference-in-Difference Analysis

	Dependent variable:				
	Rating	Sentiment	Length	Helpfulness	Image count
	(1)	(2)	(3)	(4)	(5)
Non-Amazon product	-0.113**	-0.022	6.465***	0.025***	0.012**
	(0.054)	(0.016)	(1.997)	(0.007)	(0.005)
After ban	-0.282***	-0.089***	-5.514**	0.002	0.008
	(0.064)	(0.019)	(2.398)	(0.008)	(0.006)
Non-Amazon product x After ban	0.207***	0.086***	-3.567	-0.027***	-0.021***
	(0.071)	(0.021)	(2.625)	(0.009)	(0.007)
Product Category Fixed Effects	Yes	Yes	Yes	Yes	Yes
Brand Fixed Effects	No	No	No	No	No
Observations	3,146	3,146	3,146	3,146	3,146
R^2	0.015	0.016	0.037	0.023	0.010
Adjusted R^2	-1.002	-1.000	-0.957	-0.984	-1.011
F Statistic (df = 3; 1548)	7.725***	8.253***	19.567***	12.366***	5.361***
Note:					*p*<0.1; ***p*<0.05;** p<0.01

Table 4: Difference-in-Difference Analysis with Category-Level Fixed Effects

	Dependent variable:				
	Rating	Sentiment	Length	Helpfulness	Image count
	(1)	(2)	(3)	(4)	(5)
After ban	-0.178***	-0.036**	-8.527***	-0.009	-0.005
	(0.051)	(0.015)	(2.126)	(0.007)	(0.006)
Non-Amazon product x After ban	0.123**	0.024	-1.494	-0.016*	-0.008
	(0.058)	(0.017)	(2.393)	(0.008)	(0.006)
Product Category Fixed Effects	No	No	No	No	No
Brand Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	3,146	3,146	3,146	3,146	3,146
R^2	0.007	0.003	0.040	0.018	0.010
Adjusted R^2	-0.331	-0.335	-0.286	-0.316	-0.327
F Statistic (df = 2; 2347)	8.048***	3.979**	49.271***	21.046***	11.256***
Note:					*p*<0.1; ***p*<0.05;** p<0.01

Table 5: Difference-in-Difference Analysis with Brand-Level Fixed Effects



(a) (b)

Figure 2: Coefficients and Confidence Intervals of the Interaction Terms

5.2 Robustness Checks

Following the steps described in Section 4, the estimated coefficients of the 26 weeks prior to the policy ban are all statistically zero (Figure 2), which tells us that their trends are parallel before the ban. Our future plan for additional robustness checks will be described in Section 6.

6 Conclusion and Next Steps

In summary, the results align with our hypothesis. After the ban, the nature of reviews changed: reviews became shorter with fewer images attached. However, we observe an increase in both average review rating and sentiment, which indicates that the number of incentivized reviews increased after the ban for the top-selling phone cables and tablets.

With the proof-of-concept based on our analysis on the ReviewMeta dataset, we plan to extend the analysis to include other product categories on Amazon. Before doing that, there are a few required data processing steps: (1) we need to use NLP tools to identify the stated incentivized reviews; (2) we would like to detect the unnatural reviews on the platform using the state-of-the-art ML methods.

Once the dataset is ready, we plan to extend the analysis in the following ways. First, we would explore the heterogeneous treatment effects in different sectors by adding fixed effects on product categories such as sports, clothing, and kitchenware. Second, matching would be applied to find the most similar non-Amazon products for each Amazon product. Both propensity-score matching and nearest-neighbour matching would be considered for additional robustness. Third, we will change the observation window and see if we would get consistent results. We expect the accuracy of our estimate improves as the observation windows get narrower. We also expect a decrease in the significance level of our results since the process further breaks down the sample size.

Our extension of this study would provide two contributions and implications for theory and practice. First, we identify the key characteristics of incentivized reviews. This provides us with a way to capture the changes in the number of incentivized reviews without knowing the ground truth. Second, we show that the review policy had unexpected effects on the affected parties. Our study seeks to entice practitioners to look beyond the direct effects of review policy when making platform-wide decisions.

References

Analytis, P. P., A. Delfino, J. Kämmer, M. Moussaïd, and T. Joachims (2017). Ranking with social cues: Integrating online review scores and popularity information.

Aral, S. (2014, 12). The problem with online ratings. *MIT Sloan Management Review 55*, 47–52.

Barbado, R., O. Araque, and C. A. Iglesias (2019). A framework for fake review detection in online consumer electronics retailers. *Information Processing and Management 56*(4), 1234 – 1244.

Che, Y.-K. and J. Hörner (2017, 12). Recommender Systems as Mechanisms for Social Learning*. *The Quarterly Journal of Economics 133*(2), 871–925.

Chen, P.-Y., Y. Hong, and Y. Liu (2018). The value of multidimensional rating systems: Evidence from a natural experiment and randomized experiments. *Management Science 64*(10), 4629–4647.

Chevalier, J. A. and D. Mayzlin (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research 43*(3), 345–354.

Choi, S., A. S. Mattila, H. B. V. Hoof, and D. Quadri-Felitti (2017). The role of power and incentives in inducing fake reviews in the tourism industry. *Journal of Travel Research 56*(8), 975–987.

Engler, T. H., P. Winter, and M. Schulz (2015). Understanding online product ratings: A customer satisfaction model. *Journal of Retailing and Consumer Services 27*, 113 – 120.

Gadidov, B. and J. L. Priestley (2018). *Does Yelp Matter? Analyzing (And Guide to Using) Ratings for a Quick Serve Restaurant Chain*, pp. 503–522. Cham: Springer International Publishing.

Goh, K.-Y., C.-S. Heng, and Z. Lin (2013). Social media brand community and consumer behavior: Quantifying the relative impact of user- and marketer-generated content. *Information Systems Research 24*(1), 88–107.

Heydari, A., M. Tavakoli, and N. Salim (2016). Detection of fake opinions using time series. *Expert Systems with Applications 58*, 83 – 92.

Hu, N., I. Bose, Y. Gao, and L. Liu (2011). Manipulation in digital word-of-mouth: A reality check for book reviews. *Decision Support Systems 50*(3), 627 – 635. On quantitative methods for detection of financial fraud.

Hu, N., J. Zhang, and P. A. Pavlou (2009, October). Overcoming the j-shaped distribution of product reviews. *Commun. ACM 52*(10), 144–147.

Ivanova, O. and M. Scholz (2017). How can online marketplaces reduce rating manipulation? a new approach on dynamic aggregation of online ratings. *Decision Support Systems 104*, 64 – 78.

Jindal, N. and B. Liu (2008). Opinion spam and analysis. In *Proceedings of the 2008 International Conference on Web Search and Data Mining*, WSDM '08, New York, NY, USA, pp. 219–230. Association for Computing Machinery.

KC, S. and A. Mukherjee (2016). On the temporal dynamics of opinion spamming: Case studies on yelp. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, Republic and Canton of Geneva, CHE, pp. 369–379. International World Wide Web Conferences Steering Committee.

Khern-am nuai, W., K. Kannan, and H. Ghasemkhani (2018). Extrinsic versus intrinsic rewards for contributing reviews in an online platform. *Information Systems Research 29*(4), 871–892.

Lappas, T., G. Sabnis, and G. Valkanas (2016). The impact of fake reviews on online visibility: A vulnerability assessment of the hotel industry. *Information Systems Research 27*(4), 940–961.

Lu, S. F. and H. Rui (2018). Can we trust online physician ratings? evidence from cardiac surgeons in florida. *Management Science 64*(6), 2557–2573.

Luca, M. (2011, September). Reviews, Reputation, and Revenue: The Case of Yelp.com. Harvard Business School Working Papers 12-016, Harvard Business School.

Luca, M. and G. Zervas (2015, May). Fake It Till You Make It: Reputation, Competition, and Yelp Review Fraud. Harvard business school working papers, Harvard Business School.

Mayzlin, D., Y. Dover, and J. Chevalier (2014, August). Promotional reviews: An empirical investigation of online review manipulation. *American Economic Review 104*(8), 2421–55.

Rastogi, A. and M. Mehrotra (2017). Opinion spam detection in online reviews. *Journal of Information & Knowledge Management 16*(04), 1750036.

Rayana, S. and L. Akoglu (2015). Collective opinion spam detection: Bridging review networks and metadata. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '15, New York, NY, USA, pp. 985–994. Association for Computing Machinery.

Wang, X., K. Liu, S. He, and J. Zhao (2016, November). Learning to represent review with tensor decomposition for spam detection. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Austin, Texas, pp. 866–875. Association for Computational Linguistics.

Wang, X., K. Liu, and J. Zhao (2017, July). Handling cold-start problem in review spam detection by jointly embedding texts and behaviors. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Vancouver, Canada, pp. 366–376. Association for Computational Linguistics.

Zhao, X., L. Wang, X. Guo, and R. Law (2015, 08). The influence of online reviews to online hotel booking intentions. *International Journal of Contemporary Hospitality Management 27*, 1343 – 1364.

Zhuang, M., G. Cui, and L. Peng (2018). Manufactured opinions: The effect of manipulating online product reviews. *Journal of Business Research 87*, 24 – 35.