Do Incentivized Reviews Poison the Well? Evidence from a Natural Experiment on Amazon.com

Jaecheol Park

Sauder School of Business, University of British Columbia
jaecheol.park@sauder.ubc.ca

Arslan Aziz

Sauder School of Business, University of British Columbia
arslan.aziz@sauder.ubc.ca

Gene Moo Lee

Sauder School of Business, University of British Columbia
gene.lee@sauder.ubc.ca

The rapid growth in e-commerce has led to a concomitant increase in consumers' reliance on digital word-of-mouth to inform their choices. As such, there is an increasing incentive for sellers to solicit reviews for their products. Recent studies have examined the direct effect of receiving incentives or introducing incentive policy on review writing behavior. However, since incentivized reviews are often only a small proportion of the overall reviews on a platform, it is important to understand whether their presence on the platform has spillover effects on the unincentivized reviews which are often in the majority. Using the state-of-the-art language model, Bidirectional Encoder Representations from Transformers (BERT) to identify incentivized reviews and a natural experiment caused by a policy change on Amazon.com in October 2016, we conduct the generalized synthetic control (GSC) analyses to identify the spillover effects of banning incentivized reviews on unincentivized reviews. Our results suggest that there are positive spillover effects of the ban on the review frequency and helpfulness, while there are negative spillover effects on rating, sentiment, and images, suggesting that the policy stimulates more and helpful reviews in the short-run, and negative and shorter reviews with fewer images in the long-run. Thus, we find that the presence of incentivized reviews on the platform poisons the well of reviews for frequency and helpfulness of unincentivized reviews.

Key words: incentivized reviews, spillover effect, platform design, BERT, generalized synthetic controls

1. Introduction

The rapid growth in e-commerce (eMarketer 2021) has led to a concomitant increase in consumers' reliance on digital word-of-mouth to inform their choices. The relationship between product ratings and sales has been well documented (e.g., Chevalier and Mayzlin 2006, Kwark et al. 2021, Lu et al. 2013, Wang et al. 2016, Zhu and Zhang 2010), and as such, there is an increasing incentive for sellers to solicit, and perhaps bias, the reviews for their products.

Recent studies have examined the issue of incentivized reviews and found the effect of receiving incentives or introducing incentive policy on review writing behavior (see Table 1). Such studies have deepened our understanding of the *direct* effect of incentives on the reviews. However, such incentivized reviews are often only a small proportion of the overall reviews on a platform (Chew 2016). For example, before Amazon banned incentivized reviews on their platform, the proportion of self-disclosed incentivized reviews was only 0.3% in our identification. Thus, while it is informative to understand how incentives can potentially distort such reviews, it is also important to understand how their presence on the platform has spillover effects on the unincentivized reviews which are often in the majority.

To understand the spillover effects of incentivized reviews, we identified a natural experiment caused by a policy change on Amazon.com. In October 2016, Amazon, which had up to then allowed sellers to incentivize buyers to post reviews as long as the reviews self-disclosed themselves as having been incentivized, implemented a ban on such incentivized reviews and also deleted all the past self-identified incentivized reviews posted on Amazon.com (Perez 2016).

A significant challenge in estimating the effectiveness of such a policy is the difficulty of identifying the ground-truth about which reviews are incentivized. While we do not directly observe which reviews are incentivized, we adopt the state-of-the-art language model, Bidirectional Encoder Representations from Transformers (BERT) model (Devlin et al. 2019), to infer incentivized reviews. Trained on about 4,000 human-labeled training datasets from Qiao et al. (2020)[1], our fine-tuned BERT model classifies incentivized reviews with high accuracy (Accuracy 96.31%, F1 score 96.14%, and AUC 98.55%, see Table 2 for more details).

To identify the causal effect of the ban on unincentivized reviews, we construct the control group as the products which had never received incentivized reviews and the treatment group as the products which had solicited and acquired incentivized reviews in the pre-treatment period. Based on the control and treatment groups built, we employ the generalized synthetic control (GSC) method (Xu 2017) to investigate the causal effect that this policy change has on the nature of subsequent reviews on the platform. Specifically, we examine the spillover effects of banning incentivized reviews on unincentivized reviews.

Our findings from the GSC analyses suggest that the policy banning incentivized reviews caused products to receive more reviews that are more helpful, but receive more negative and shorter reviews with fewer images. Additional analyses find that the effect on review frequency and helpfulness diminishes in the long-run while the effects on the rating, sentiment, length, and images persist for an extended time period.

Table 1   Summary of Empirical Studies on Monetary Incentives in Online Review Context

		Focus of the Study	
		Reviewer (Receiving Incentives)	Platform (Changing Incentive Policy)
Type of the Effect	Direct	Khern-am-nuai et al. (2018) (Sec 5.2.1), Yu et al. (2022)	Wang et al. (2012), Wang et al. (2016), Burtch et al. (2018), Khern-am-nuai et al. (2018)
	Spillover	Qiao et al. (2020)	**Our study**

2.   Related Literature

Extant literature on the impact of monetary incentives on online reviews can be summarized into four quadrants using two dimensions as shown in Table 1. First, some studies have investigated the *direct* effect of receiving monetary incentives on individual reviewers (upper left quadrant). Reviewers who received monetary incentives posted fewer reviews without changing their review length (Khern-am-nuai et al. 2018). However, when reviewers received performance-contingent monetary incentives, they increased the number and length of reviews and wrote more helpful reviews (Yu et al. 2022).

Second, many studies have examined the *direct* effect of changing the monetary incentive policy on the platform level (upper right quadrant in Table 1). Introducing completion-contingent incentives had no effect on the review helpfulness, but providing reviewers performance-contingent incentives and requiring them to disclose sponsorship increased the review helpfulness and length of reviews (Wang et al. 2012). Also, introducing the policy to allow incentives to reviewers increased review volume, but the reducing monetary incentives for writing reviews decreased review volume and product sales (Wang et al. 2016). Similarly, monetary incentives induced more reviews, but did not increase the length of reviews, however informing reviewers of the number of reviews written by other customers increased the number of reviews and review length (Burtch et al. 2018). Finally, introducing incentives resulted in more positive reviews, but decreased the length and helpfulness of reviews (Khern-am-nuai et al. 2018).

Thus, while our understanding of the *direct* effect of incentives on the reviewers or the platform has matured, as yet little attention has been paid to the *spillover* effects of such incentives on unincentivized reviews and reviewers. One study in this area focuses on examining the effects of receiving incentives on the same reviewer's subsequent unincentivized reviews (Qiao et al. 2020) (bottom left quadrant in Table 1). The study finds that receiving monetary incentives had spillover

---

[1] We thank Dandan Qiao for generously providing us with the data from Qiao et al. (2020).

4

Park, Aziz, and Lee: *Do Incentivized Reviews Poison the Well?*
Article submitted to *BU Platform Strategy Research Symposium*; manuscript no. 50

effects on incentivized reviewers' subsequent unincentivized reviews, which were more frequent, positive, shorter, and of less linguistic effort.

Our study broadens the scope of *spillover* effects (1) by analyzing the entire unincentivized reviewer population who are often in the majority on most popular review platforms and (2) by examining the ban of incentivized reviews, which may not be symmetric with the introduction of incentivized reviews. Thus, understanding how the extinction of a small proportion of incentivized reviews may have spillover effects on the majority of unincentivized reviews is an important question that we answer in this study.

3. Data and Method
3.1. Institutional Settings

Prior to October 2016, Amazon used to allow sellers to provide incentives to customers in the form of free samples or discounted products in exchange for honest reviews (Burtch et al. 2018, Qiao et al. 2020). Following Federal Trade Commission (FTC) guidelines, Amazon set strict Community Guidelines[2] for such incentivized reviews, requiring reviewers explicitly self-disclose the fact that they have received a free product or a discount in the form of a statement such as "I received this product either free or at a discount rate in exchange for my honest and unbiased review" embedded in the review text, as shown in Figure 1.

Originally, Amazon allowed these incentivized reviews, anticipating positive effects in that reviewers could write an honest, unbiased, and true opinion about the product and that incentivized reviews may help new sellers to increase their reputation or recognition on the crowded platform, i.e., to address the cold-start problem (Chew 2016). However, incentivized reviews are well-documented as being biased in favor of the seller offering the incentive (see Table 1). Therefore, Amazon decided to prohibit incentivized reviews on October 3, 2016 (Chew 2016, Perez 2016).

---

[2] https://www.amazon.com/gp/help/customer/display.html?nodeId=201602680 (Accessed May 20, 2022)

Dawn Michelle
★★★★☆ **Really nice**
Reviewed in the United States on July 17, 2016
Verified Purchase

I always want the latest gadgets so I looked at action cameras and I couldn't see paying the prices out there for GoPro and if my grandkids come over and they want to run around the ranch and play with it, I won't have to be concerned that it cost way to much for them to touch. I used it around the house and in our pond and it stayed dry the whole time it was in the pond. The pics were just as good as any other action cameras. I gave it 4 starts because it came with only a out of country wall charger. I also had a hard time opening the case. But I am sure it was user error.

Disclaimer: I received this product either free or at a discounted rate in exchange for my honest and unbiased review, which I have given
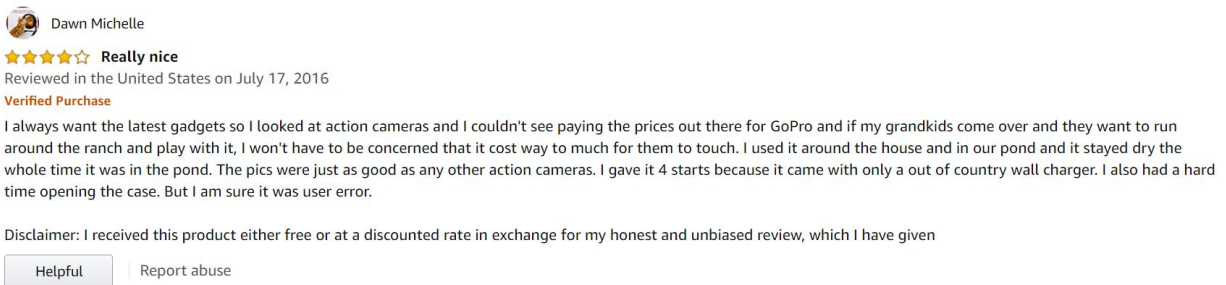
Helpful   Report abuse

Figure 1   An Example of Incentivized Review on Amazon.com

3.2. Data

Our primary data sources consist of Amazon review data and incentivized review data. Amazon review data is publicly available for academic research and contains over 200 million reviews from May 1996 to October 2018 (Ni et al. 2019). Since our focus is on identifying the impact of banning incentivized reviews, we select reviews posted from 8 weeks prior to, and until 24 weeks after, the policy implementation on October 3, 2016. The data includes review-related information such as ratings, text, and helpfulness votes and product-related information such as product descriptions, price, and category.

We also used incentivized review data used to train the machine learning model in Qiao et al. (2020) to help address a key operationalizational challenge, the identification of incentivized reviews among the entire review sample. The data contains about 4,000 manually-labeled reviews, among which half are incentivized reviews and the other half are unincentivized reviews (see Online Appendix II in Qiao et al. (2020)), and is used to fine-tune our machine learning model to identify the incentivized reviews. To examine the spillover effects of banning incentivized reviews, we go through three steps as illustrated in Figure 2.

3.3. Step 1: Identification of Incentivized Reviews

First, for identifying incentivized reviews, we use the state-of-the-art language model, BERT, which is known as one of the best performing models on various natural language processing tasks (Devlin et al. 2019). Applying the BERT model consists of two parts: pre-training and fine-tuning as shown in Figure 3. During pre-training done by Google, developers used 16 GB of text data including 800 million words from the Books corpus and 2,500 million words from the English Wikipedia for
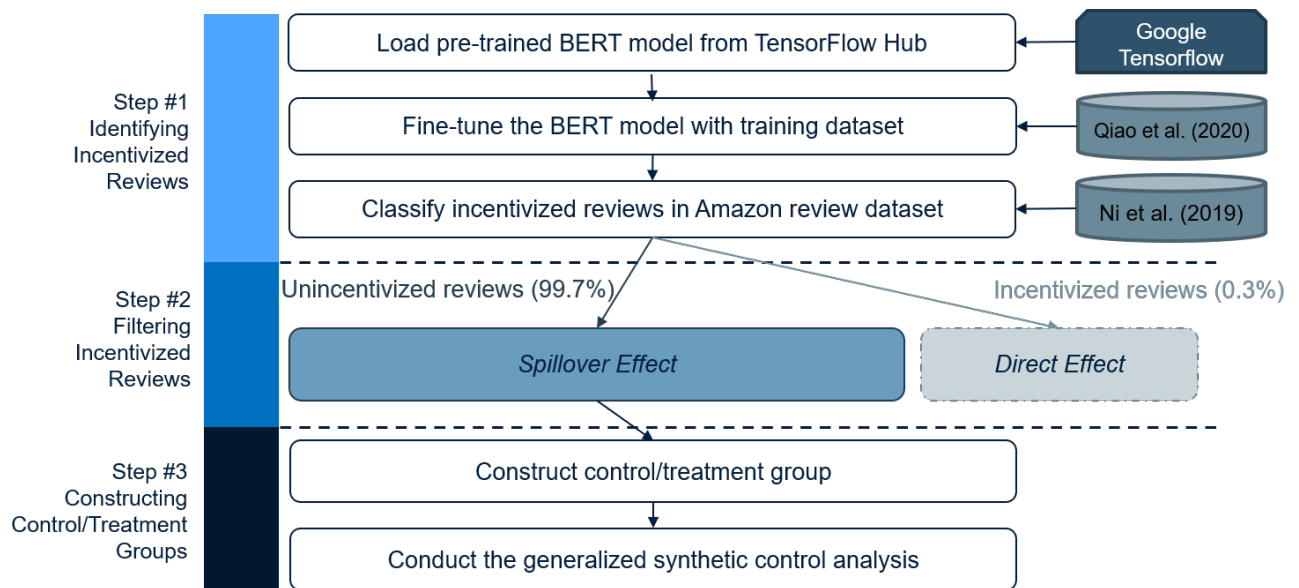

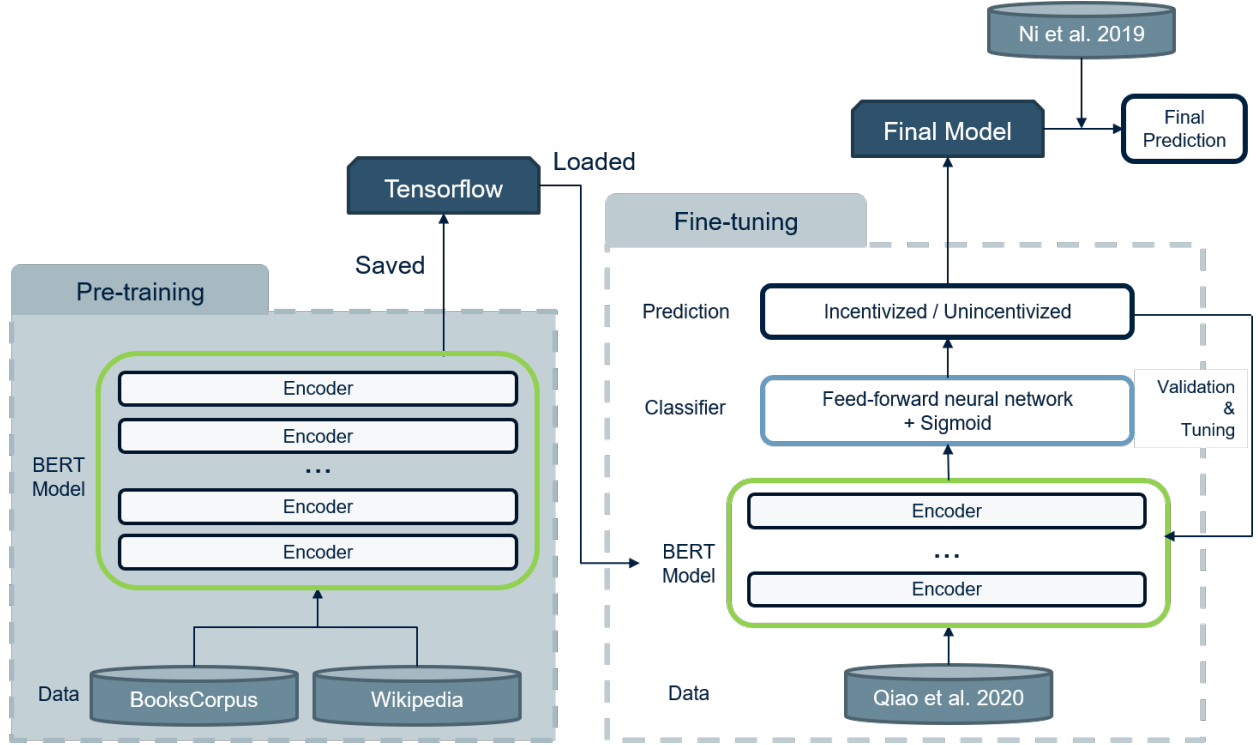**Figure 2   Our Study's Research Framework**

Figure 3   Pre-training and Fine-tuning of BERT

various BERT models with different architectures, and saved these pre-trained models on the model hub. Among various BERT model architectures, we adopted the BERT LARGE model which has 24 layers, 1024 hidden nodes, and 340M parameters for our research.

During the fine-tuning step we performed, the BERT model is first initialized with the pre-trained parameters, and these parameters are fine-tuned with our manually-labeled incentivized review data from Qiao et al. (2020). We split the data into training, validation, and test sets (72%, 18%, and 10% respectively), iteratively validating and fine-tuning the model to achieve the best performance. Cross-validation results in Table 2 show that our fine-tuned model can classify incentivized reviews with high accuracy (Accuracy 96.31%, F1 score 96.14%, and AUC 98.55%).

With the final trained model, we classify the incentivized reviews among the entire Amazon review sample. Considering that our training dataset is sentence-level, we split each Amazon review into sentences, predict whether each sentence is incentivized or not, and aggregate it into review-level. Our identification shows that before the banning the incentivized reviews, there were 7,082 incentivized reviews out of 2,367,103 reviews, showing that the proportion of self-disclosed incentivized reviews was only 0.3% aligning with the announcement from Amazon.com that "incentivized reviews make up only a tiny fraction of the tens of millions of reviews on Amazon" (Chew 2016).

Table 2   **Performance of BERT Model**

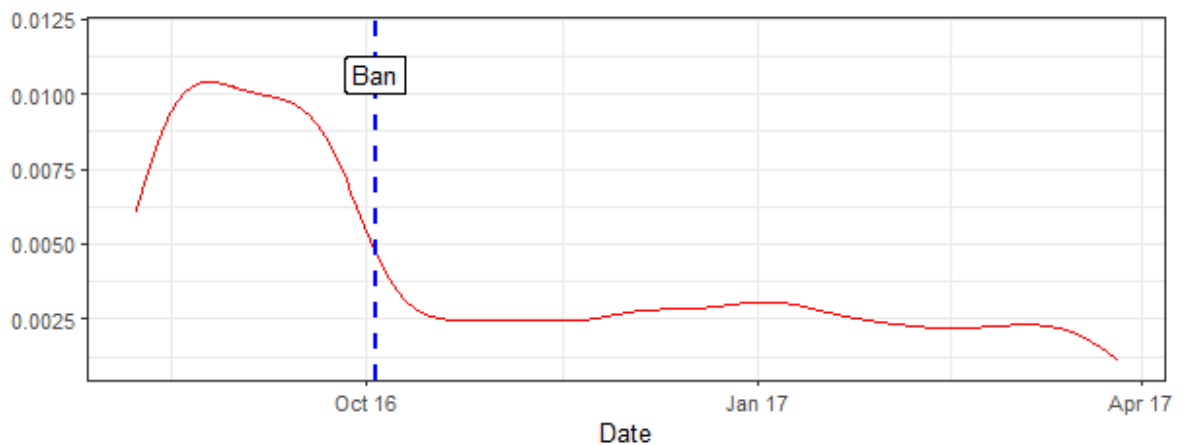| Epochs	Accuracy	Precision	Recall	F1 Score	AUC
1	0.9570	0.9751	0.9362	0.9552	0.9889
2	0.9583	0.9624	0.9521	0.9572	0.9794
3	0.9583	0.9751	0.9388	0.9566	0.9758
4	0.9622	0.9651	0.9574	0.9613	0.9790
Testing	0.9631	0.9444	0.9791	0.9614	0.9855

Figure 4   Trend of Incentivized Reviews Proportions in Our Sample Period

3.4.   Step 2: Filtration of Incentivized Reviews

In the second step, given the identification of incentivized reviews, we filter them out to investigate the spillover effects of banning incentivized reviews on unincentivized reviews. Figure 4 represents the distribution of incentivized reviews across the time period and shows that incentivized reviews are heavily identified before the policy ban as we expected. Specifically, we filtered 13,688 incentivized reviews identified in the first step out of 11,135,817 reviews in our entire sample period.

3.5.   Step 3: Construction of Control and Treatment Groups

Finally, we construct control and treatment groups to estimate the spillover effects of banning incentivized reviews on unincentivized reviews using GSC approach. We build a control group with unincentivized products and a treatment group with incentivized products which are also viewed by the customers who viewed the products in the control group. We use unincentivized products as the control group because the sellers did not provide incentives to buyers to write reviews or recruit incentivized reviewers from third-party platforms in the pre-treatment period and hence, they are unaffected by the policy banning incentivized reviews. The co-viewed incentivized products, on the

8

Park, Aziz, and Lee: *Do Incentivized Reviews Poison the Well?*
Article submitted to *BU Platform Strategy Research Symposium*; manuscript no. 50

Table 3   Summary Statistics of Control and Treatment Groups

Variable	Group	N	Mean	St. Dev	Min	Med	Max
Rating	Control	18,879	4.22	1.05	1	4.67	5
	Treatment	2,750	4.21	1.14	1	5	5
Sentiment	Control	18,879	0.48	0.36	-1	0.568	1
	Treatment	2,750	0.52	0.39	-0.99	0.625	1
Length	Control	18,879	35.10	45.6	1	23.7	1,957
	Treatment	2,750	40.00	56.2	1	24	2,088
Images	Control	18,879	0.02	0.256	0	0	20
	Treatment	2,750	0.06	0.457	0	0	23
Frequency	Control	18,879	4.06	7.97	1	2	465
	Treatment	2,750	2.11	2.46	1	1	73
Helpfulness	Control	18,879	0.81	4.87	0	0	577
	Treatment	2,750	0.89	3.87	0	0	208

Note. Summary statistics are product-week level.

other hand, can be affected by the policy because the reviews for incentivized products consist of both incentivized and unincentivized ones.

We define an *co-viewed incentivized product* as a product if (1) 10% or more of its reviews in the pre-treatment period are classified as *incentivized reviews* and (2) *co-viewed* with the focal unincentivized products. Specifically, we identified the incentivized products in the list of co-viewed products from the unincentivized products or vice versa to find closely related control and treatment groups. Additionally, we restricted both control (unincentivized) and treatment (incentivized) groups to products which receive at least 10 reviews in 6 months before the ban.

Our final control group has 18,879 unincentivized products and the treatment group has 2,750 co-viewed incentivized products. Table 3 presents the comparison of weekly-level descriptive statistics of dependent variables between the control and treatment groups. The first week of the post-treatment period data was dropped to estimate the effect of policy implementation accurately.

4.   Hypotheses Development

In this section, we develop hypotheses regarding the spillover effects of banning monetary incentives on unincentivized reviews for previously incentivized products based on four review measurements and two possible directions. First, we adopt the four widely used measurements of reviews: review valence, reviewers' effort, review quantity, and review quality. Second, banning incentivized reviews can affect subsequent unincentivized reviews in two opposite directions: positive or negative spillover effects. Finally, incentivized products have two types of reviewers: incentivized or unincentivized reviewers.

4.1. Effect on Review Valence and Reviewers' Effort

Review valance measures suggest reviewers' experience of products (Khern-am-nuai et al. 2018). We use *rating* (Chintagunta et al. 2010, Khern-am-nuai et al. 2018, Qiao et al. 2020), and *sentiment* of review (Khern-am-nuai et al. 2018, Qiao et al. 2020) as review valence measures in our study. Similarly, we adopt review *length*, i.e., word count (Mudambi and Schuff 2010, Khern-am-nuai et al. 2018, Qiao et al. 2020), and *images*, i.e., the number of images as reviewers' level of effort.

Based on the discussion of the extant literature on the effect of monetary incentives on online reviews in Section 2, due to the negative effect of monetary incentives on intrinsic motivation, incentivized reviewers write short unincentivized reviews (Khern-am-nuai et al. 2018, Qiao et al. 2020). Therefore, we posit that banning monetary incentives increases intrinsic motivation, and increased motivation can lead to a higher level of effort from incentivized reviewers.

Regarding unincentivized reviewers, we can expect similar outcomes. As incentivized reviews include an explicit statement of disclosing their sponsorship from sellers as presented in Figure 1, unincentivized reviewers can easily identify the marketing intent of reviews. In this case, unincentivized reviewers can question the credibility and objectivity of reviews, raise mistrust on the platform (Boush et al. 1994, Campbell and Kirmani 2000, Friestad and Wright 1995), and disregard the positive information in the reviews (Awad and Ragowsky 2008, Brown and Krishna 2004). As a result, with the presence of incentivized reviews, customers will take caution during information processing (Kramer 1998, Nickerson 1998). However, if the monetary incentives are banned and all incentivized reviews are deleted, we can expect that customers recover trust and become less conservative, leading to a higher level of effort into review writing behavior and increasing the valence of their reviews.

Hypothesis 1a*: Review valence (measured by rating and sentiment) and reviewers' effort (measured by length and images) for the previously incentivized products will increase after banning incentivized reviews.*

However, it is possible that banning incentivized reviews may cause the opposite effect. Incentivized reviewers write more positive unincentivized reviews due to an elevated feeling of being controlled and a receded feeling of autonomy (Qiao et al. 2020) or due to the reciprocity affecting reviewers' transformation process between the experience of a product and the review writing behavior (Khern-am-nuai et al. 2018). Therefore, we postulate that banning monetary incentives increases the feeling of autonomy, decreases the feeling of being controlled, or alleviates the feeling of reciprocity. As a result, incentivized reviewers decrease the valence of reviews after the ban.

Also, reviewers who no longer receive the incentives after the ban can lose motivation to write reviews, making them put less effort to write reviews.

Similarly, unincentivized reviewers decrease the valence. Extant literature documents that ratings show a declining trend (Li and Hitt 2013, Wu and Huberman 2008, Moon et al. 2010, Moe and Trusov 2011, Godes and Silva 2012) in general. Therefore, in the absence of incentivized reviews after the ban, we can expect that unincentivized reviewers decrease the valence of reviews and their effort.

Hypothesis 1b*: Review valence (measured by rating and sentiment) and reviewers' effort (measured by length and images) for the previously incentivized products will decrease after banning incentivized reviews.*

4.2. Effect on Review Quantity

Review quantity measure implies the reviewer's engagement (Qiao et al. 2020) and we use *frequency*, i.e., the number of reviews (Qiao et al. 2020), as a review quantity measure in our study. In the same vein as above, we derive two competing hypotheses for review quantity.

Based on the findings from the literature in Section 2, reviewers write fewer unincentivized reviews (Qiao et al. 2020) and overall reviews (Khern-am-nuai et al. 2018) after receiving incentives. In addition, the substitution effect, which incentivized reviewers write unincentivized reviews instead of incentivized reviews, may occur because of their difficulty in writing incentivized reviews after the ban. Therefore, we expect that the ban on incentivized reviews on the platform induces more reviews.

Hypothesis 2a*: Review quantity (measured by frequency) for the incentivized products will increase after banning monetary incentives.*

However, we can also present the opposite hypothesis for banning incentivized reviews on review quantity. Reviewers write more reviews after introducing a monetary incentive policy on the platform (Burtch et al. 2018) and write fewer reviews after reducing monetary incentives for writing reviews (Wang et al. 2016). Also, reviewers who lose their monetary incentives after the ban can be less motivated to write reviews, making them write fewer reviews. Therefore, we can posit that banning monetary incentives for writing reviews decreases the frequency of reviews.

Hypothesis 2b*: Review quantity (measured by frequency) for the incentivized products will decrease after banning monetary incentives.*

Park, Aziz, and Lee: *Do Incentivized Reviews Poison the Well?*
Article submitted to *BU Platform Strategy Research Symposium*; manuscript no. 50

11

4.3. Effect on Review Quality

The *helpfulness* of reviews presents "the diagnostic value of review for decision-making" (pp. 1059, Yin et al. 2021) and is measured by helpful votes (Ghose and Ipeirotis 2011, Khern-am-nuai et al. 2018) as a review quality measure in our study.

Review helpfulness is contingent on other review characteristics such as rating (Mudambi and Schuff 2010, Chua and Banerjee 2016, Eslami et al. 2018), sentiment (Schindler and Bickart 2012, Salehan and Kim 2016, Yin et al. 2016), length (Mudambi and Schuff 2010, Ghose and Ipeirotis 2011, Schindler and Bickart 2012, Eslami et al. 2018), product type (Chua and Banerjee 2016), information quality (Chua and Banerjee 2016), product information (Schindler and Bickart 2012), emotion (Yin et al. 2014, 2021), and so on.

Moreover, existing literature on monetary incentives in online reviews presented in Table 1 shows the mixed findings on the effect of receiving monetary incentives or introducing incentive policy generally on review helpfulness: no effect (Wang et al. 2012, Qiao et al. 2020), positive effect when performance-contingent incentives are provided (Wang et al. 2012, Yu et al. 2022), and negative effect (Khern-am-nuai et al. 2018). Considering our competing hypotheses developed for *rating, sentiment, length, and images*, and the mixed findings from literature, we derive the last two hypotheses on review helpfulness.

Hypothesis 3a: *Review quality (measured by helpfulness) for the incentivized products will increase after banning monetary incentives.*

Hypothesis 3b: *Review quality (measured by helpfulness) for the incentivized products will decrease after banning monetary incentives.*

5. Empirical Analyses and Results
5.1. Generalized Synthetic Control Methods

We employ the generalized synthetic control (GSC) method (Xu 2017) to examine the spillover effects of banning incentivized reviews on the nature of subsequent unincentivized reviews on the platform. GSC method integrates the interactive fixed effects (IFE) model (Bai 2009) and the synthetic control (SC) method (Abadie et al. 2010, 2015) and is widely adopted in recent studies from information systems (He et al. 2020, Pattabhiramaiah et al. 2021, Wang et al. 2021, Chen et al. 2022) and marketing (Guo et al. 2020, Puranam et al. 2021). GSC method has several advantages: (1) it allows a large number of treated units, which are incentivized products in our study, (2) it generates synthetic control units from multiple control units by matching outcomes in the pre-treatment period like SC method, and (3) it includes the interactive fixed effects to alleviate the time-varying confounders like IFE model.

We perform the GSC analysis with the following specification:

$$DV_{it} = \delta_{it} Treat_i After_t + X_{it}\beta + \lambda'_i f_t + \epsilon_{it} \tag{1}$$

where DV_{it} represents the dependent variable of interest including *Rating*, *Sentiment*, *Length*, *Images*, *Frequency*, and *Helpfulness* which are the weekly-level mean rating of reviews, mean sentiment of reviews, mean number of the word in reviews, mean number of images posted in reviews, number of reviews, and mean number of helpful votes for reviews for each product i and week t, respectively. $After_t$ is a dummy variable equal to 1 if the review is written after the policy ban and 0 otherwise, $Treat_i$ is a dummy variable equal to 1 if the review is written for products incentivized in the pre-treatment period and 0 otherwise. The coefficient δ_{it} for the interaction term is the parameter of interests, which represents the spillover effects of the policy ban on the dependent variables.

In addition, X_{it} is a vector of observed covariates from product information such as price, category, and brand and β is a vector of the corresponding estimates. f_t is a vector of unobserved common factors and λ'_i is a vector of unknown factor loadings, which are obtained by the cross-validation performance of pre-treatment period fit (see Xu (2017) for more details). Note that $\lambda'_i f_t = \lambda_{i1} f_{1t} + \lambda_{i2} f_{2t} + ... + \lambda_{ir} f_{rt}$ and, when r=2, $f_{1t} = 1$, and $\lambda_{i2} = 1$, $\lambda'_i f_t$ becomes $\lambda_{i1} + f_{2t}$ where λ_{i1} representing the product fixed effect and f_{2t} representing the week fixed effect are the special cases of a two-way fixed-effects specification. ϵ_{it} represents the error term.

5.2. Main Results

The results of the GSC analysis for each dependent variable are shown in Table 4. The coefficients of the interaction terms for the *Sentiment*, *Length*, and *Images* are negative and significant (Columns 2, 3, 4), suggesting that after banning incentivized reviews, previously incentivized products receive lower review valence and less effort by the reviewers (i.e., shorter, fewer images). However, the

Table 4   GSC Estimation Results for 8 Weeks Prior and After Ban

Dependent Variables	(1) *Rating*	(2) *Sentiment*	(3) *Length*	(4) *Images*	(5) *Frequency*	(6) *Helpfulness*
Treat × *After*	-0.0353	-0.0322***	-11.0696***	-0.0420***	0.4211**	0.2714*
	(0.0309)	(0.0119)	(1.4759)	(0.0076)	(0.1673)	(0.1560)
Week Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Treatments	319	319	319	319	319	319
Number of Controls	8,276	8,276	8,276	8,276	8,276	8,276

Note. ***p < 0.01; **p < 0.05; *p < 0.1.

	Table 5	**Summary of Results of the Hypothesis Testing**	
Hypothesis	Review Measures	Measurements	Results
1	Review Valence	*Rating*	H1 not supported
		Sentiment	H1b supported
	Reviewer's Effort	*Length*	H1b supported
		Images	H1b supported
2	Review Quantity	*Frequency*	H2a supported
3	Review Quality	*Helpfulness*	H3a supported

coefficients of the interaction terms for the *Frequency* and *Helpfulness* are positive and significant (Columns 5, 6), suggesting that the policy caused incentivized products to receive more and helpful unincentivized reviews. Finally, there are no significant coefficients for *Rating* (Column 1). Overall, we can find that the negative spillover effects on valence and effort of unincentivized reviews for incentivized products, while the positive spillover effects of banning incentivized reviews on quantity and quality of unincentivized reviews.

Moreover, the estimates show that the treatment effect is economically meaningful. The negative coefficients of the interaction terms show that the policy caused the 4% decrease in the sentiment, the 28% decrease in review length, and the 72% decrease in the number of attached images, while the positive coefficients of the interaction terms show that the policy caused the 20% increase in the review count and 30% increase in the helpfulness score of unincentivized reviews for incentivized products. Table 5 summarizes the results of hypothesis testing.

5.3. Temporal Time Trends after Treatment

We further investigate the temporal nature of the treatment effect of the policy banning incentivized reviews. In the review manipulation context, soliciting fake reviews leads to a significant increase in average rating and sales rank but the effect is short-lived (Proserpio et al. 2020). Moreover, all types of review manipulation have short-term benefits, but excessive manipulation leaves cues of alteration and raises suspicions, leading to a negative effect on product performance (Zhuang et al. 2018). Similarly, a sequential effect of online reviews has also been found. Subsequent negative feedback ratings arrive more rapidly than the first one (Cabral and Hortacsu 2010) and the decrease in rating remains after controlling for time, reviewer, and product effects (Godes and Silva 2012).

To distinguish the short- and long-term effects of the policy, we replicate our analysis with different periods: 16 weeks and 24 weeks after the policy implementation. The results of GSC analyses for each dependent variable are shown in Table 6. The coefficients of the interaction term for *Rating*, *Sentiment*, *Length*, and *Images* are negative and significant across the two panels (Columns 1, 2, 3, 4), suggesting that the effects persist for an extended time period after the ban.

14

Park, Aziz, and Lee: *Do Incentivized Reviews Poison the Well?*
Article submitted to *BU Platform Strategy Research Symposium*; manuscript no. 50

Table 6 GSC Estimation with Different Time Periods

Dependent Variables	(1) Rating	(2) Sentiment	(3) Length	(4) Images	(5) Frequency	(6) Helpfulness
Panel A: 8 Weeks Before and 16 Weeks After Ban						
Treat × After	-0.0771***	-0.0398***	-12.1047***	-0.0504***	0.1596	0.2609*
	(0.0244)	(0.0090)	(1.2999)	(0.0071)	(0.1856)	(0.1381)
Week Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Treatments	323	323	323	323	323	323
Number of Controls	10,696	10,696	10,696	10,696	10,696	10,696
Panel B: 8 Weeks Before and 24 Weeks After Ban						
Treat × After	-0.0678***	-0.0426***	-12.7857***	-0.0513***	0.2274	0.2489
	(0.0227)	(0.0082)	(1.1034)	(0.0058)	(0.1710)	(0.2307)
Week Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Product Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Number of Treatments	323	323	323	323	323	323
Number of Controls	12,135	12,135	12,135	12,135	12,135	12,135

Note. ***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

Notably, the coefficient of the interaction term for *Frequency* is insignificant across two panels (Column 5) and the interaction term for *Helpfulness* becomes insignificant in the second panel (Column 6), suggesting that the effect did not last long. Also, we plot the estimated average treatment effect on the treated and the counterfactual for each dependent variable in Figure 5.

6. Conclusion
6.1. Theoretical Contributions
Our study contributes to the literature on the impact of monetary incentives on online product reviews. First, our study broadens the scope of the effects of monetary incentives on online reviews. Previous literature investigated the direct effect of monetary incentives on online reviews by focusing on the introduction of an incentive policy (Wang et al. 2012, 2016, Burtch et al. 2018, Khern-am-nuai et al. 2018, Yu et al. 2022) or the spillover effect within incentivized reviewers on their subsequent unincentivized review characteristics (Qiao et al. 2020). We expand the scope by analyzing the spillover effect of banning incentivized reviews, which is not a symmetric effect of introducing incentivized reviews, on all unincentivized reviews which account for the majority of review platforms. We find empirical evidence that, even when present in small numbers, incentivized reviews 'poison the well' of reviews, and that banning them stimulates more unincentivized reviews that are more helpful.

Second, we add the results to the literature that review manipulation has short-term effects (Proserpio et al. 2020, Zhuang et al. 2018), by finding that the effects of banning incentivized
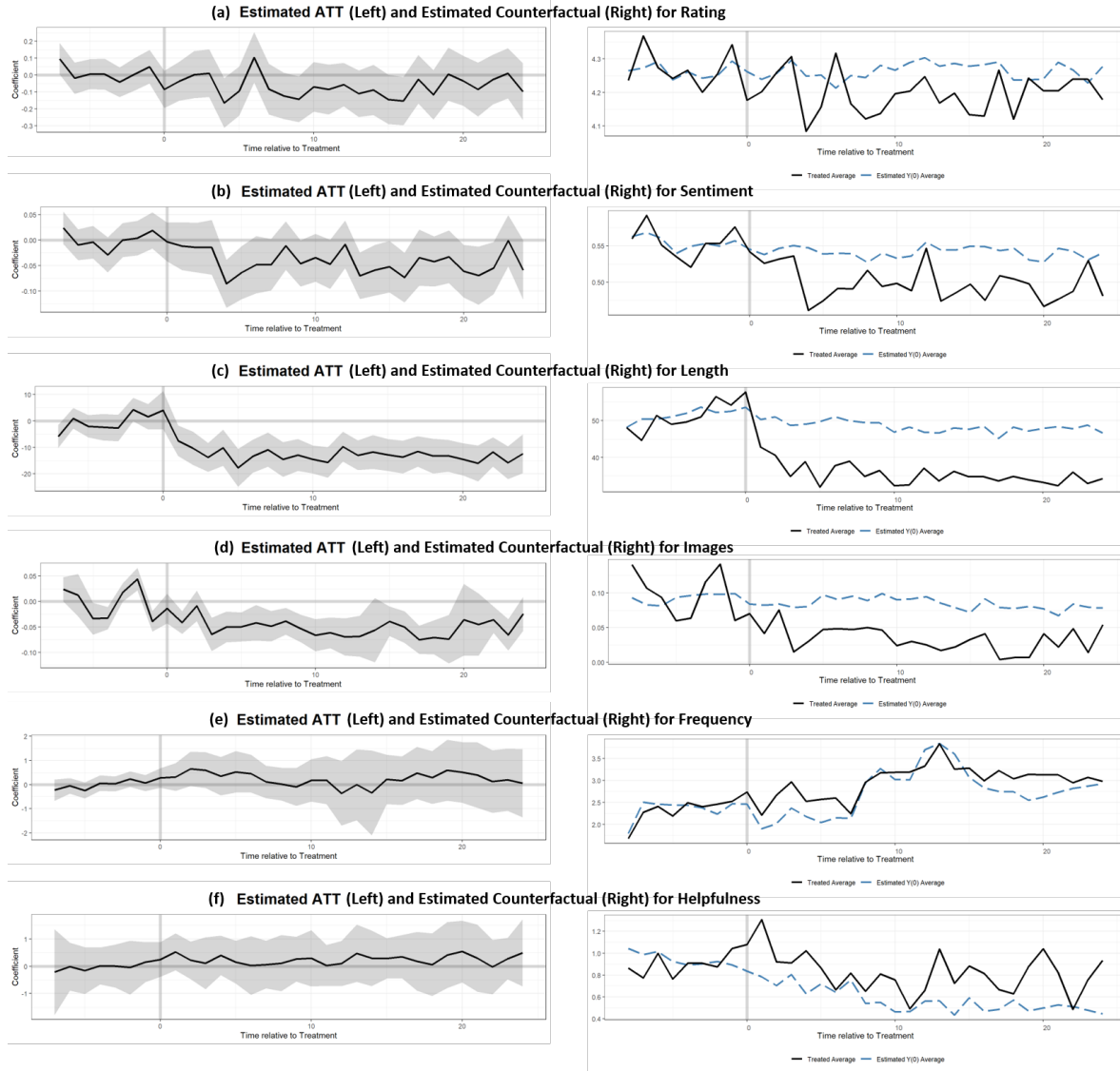
Figure 5   Estimated Average Treatment Effect on the Treated (ATT) and Counterfactual Plots.

reviews differ in the short vs. long-run. The positive effects on review frequency and helpfulness diminish shortly while the negative effects on review rating, sentiment, length, and images last for an extended time period.

6.2.   Managerial Implications

Our study also yields important managerial implications regarding the trust of users in the review platform. Monetary incentive policies for reviews have become prevalent on e-commerce websites and review platforms to attract more reviews to enrich their ecosystem or to solve the cold-start problem on their platform. However, platform managers should be alert to the possibility of the negative effect of introducing an incentive policy where allowing incentivized reviews on the platform

can lower overall trust in the review platform, which in turn can cause lower effort by unincentivized reviewers to write more informative and helpful reviews. Our empirical evidence suggests that incentivized reviews 'poison the well' of reviews.

Our results can also provide guidance to platform managers who have already introduced the incentive policy and are suffering from low-quality, uninformative, and unhelpful reviews on their platform by showing the positive spillover effects of one simple strategy the platform can implement to increase the frequent and helpful reviews: to ban incentivized reviews.

6.3. Limitations and Future Directions

We acknowledge the limitations of this study. First, because of the lack of sales data, we cannot estimate the effect of the policy banning incentivized reviews on product sales. Second, future research can probe how the policy change affects individual reviewers' psychological status such as trust or review writing behavior with lab experiments to more fully explore the mechanism driving the observed phenomena. Finally, our study does not differentiate effects on product types (e.g., niche vs. popular products; search vs. experience goods), as well as on reviewer types such as incentivized or unincentivized reviewers. These limitations of our study provide promising avenues for further research.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of American Statistical Association* **105**(490) 493–505.

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2015. Comparative Politics and the Synthetic Control Method. *American Journal of Political Science* **59**(2) 495–510.

Awad, Neveen F. and Arik Ragowsky. 2008. Establishing Trust in Electronic Commerce Through Online Word of Mouth: An Examination Across Genders. *Journal of Management Information Systems* **24**(4) 101-121.

Bai, Jushan . 2009. Panel Data Models with Interactive Fixed Effects. *Econometrica* **77**(4) 1229–1279.

Boush, David M., Marian Friestad, and Gregory M. Rose (1994) Adolescent Skepticism toward TV Advertising and Knowledge of Advertiser Tactics. *Journal of Consumer Research* **21**(1) 165–175.

Brown, Christina L. and Aradhna Krishna. 2004. The Skeptical Shopper: A Metacognitive Account for the Effects of Default Options on Choice. *Journal of Consumer Research* **31**(3) 529–539.

Burtch, Gordon, Yili Hong, Ravi Bapna, and Vladas Griskevicius. 2018. Stimulating Online Reviews by Combining Financial Incentives and Social Norms. *Management Science* **64**(5) 2065–2082.

Cabral, Luis, and Ali Hortacsu. 2010. The Dynamics of Seller Reputation: Evidence from eBay. *The Journal of Industrial Economics* **58**(1) 54-78.

Campbell Margaret C. and Amna Kirmani (2000) Consumers' Use of Persuasion Knowledge: The Effects of Accessibility and Cognitive Capacity on Perceptions of an Influence Agent. *Journal of Consumer Research* **27**(1) 69–83.

Chen, Wei, Zaiyan Wei, and Karen Xie. 2022. The Battle for Homes: How Does Home Sharing Disrupt Local Residential Markets?. *Management Science* **0**(0).

Chevalier, Judith A. and Dina Mayzlin. 2006. The Effect of Word of Mouth on Sales: Online Book Reviews. *Journal of Marketing Research* **43**(August) 345–354.

Chew, Chee. 2016. Update on Customer Reviews. Accessed May 23, 2022, https://www.aboutamazon.com/news/innovation-at-amazon/update-on-customer-reviews.

Chintagunta, Pradeep K. , Shyam Gopinath, and Sriram Venkataraman .2010. The Effects of Online User Reviews on Movie Box Office Performance: Accounting for Sequential Rollout and Aggregation Across Local Markets. *Marketing Science* **29**(5) 944-957.

Chua, Alton Y.K. and Snehasish Banerjee. 2016. Helpfulness of User-Generated Reviews as a Function of Review Sentiment, Product Type and Information Quality. *Computers in Human Behavior* **54** 547-554.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding. *arXiv* 1810.04805.

Eslami, Seyed Pouyan, Maryam Ghasemaghaei, and Khaled Hassanein. 2018. Which Online Reviews Do Consumers Find Most Helpful? A Multi-Method Investigation. *Decision Support Systems* **113** 32–42.

Friestad, Marian and Peter Wright. 1995. Persuasion Knowledge: Lay People's and Researchers' Beliefs about the Psychology of Advertising. *Journal of Consumers Research* **22**(1) 62–74.

eMarketer. 2021. Worldwide eCommerce Will Approach \$5 Trillion This Year. Accessed May 23, 2022, https://www.emarketer.com/content/worldwide-ecommerce-will-approach-5-trillion-this-year.

Ghose, Anindya and Panagiotis G. Ipeirotis. 2011. Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics. *IEEE Transaction on Knowledge and Data Engineering* **23**(10) 1498-1512.

Godes, David and Jose C. Silva. 2012. Sequential and Temporal Dynamics of Online Opinion. *Marketing Science* **31**(3) 448-473.

Guo, Tong, Srinivasaraghavan Sriram, and Puneet Manchanda. 2020. "Let the Sunshine In": The Impact of Industry Payment Disclosure on Physician Prescription Behavior. *Marketing Science* **39**(3) 516-539.

He, Shu, Jing Peng, Jianbin Li, and Liping Xu. 2020. Impact of Platform Owner's Entry on Third-Party Stores. *Information Systems Research* **31**(4) 1467-1484.

Khern-am-nuai, Warut, Karthik Kannan, and Hossein Ghasemkhani. 2018. Extrinsic versus Intrinsic Rewards for Contributing Reviews in an Online Platform. *Information Systems Research* **29**(4) 871-892.

Kramer, Roderick M. 1998. Paranoid Cognition in Social Systems: Thinking and Acting in the Shadow of Doubt. *Personality and Social Psychology Review* **2**(4) 251–275.

Kwark, Young, Gene Moo Lee, Paul A. Pavlou, and Liangfei Qiu. 2021. On the Spillover Effects of Online Product Reviews on Purchases: Evidence from Clickstream Data. *Information Systems Research* **32**(3) 895-913.

Li, Xinxin and Lorin M. Hitt. 2008. Self-Selection and Information Role of Online Product Reviews. *Information Systems Research* **19**(4) 456-474.

Lu, Xianghua, Sulin Ba, Lihua Huang, and Yue Feng. 2013. Promotional Marketing or Word-of-Mouth? Evidence from Online Restaurant Reviews. *Information Systems Research* **24**(3) 596-612.

Moe, Wendy W. and Michael Trusov. 2011. The Value of Social Dynamics in Online Product Ratings Forum. *Journal of Marketing Research* **48**(3) 444–456.

Moon, Sangkil, Paul K. Bergey, and Dawn Iacobucci. 2010. Dynamic Effects among Movie Ratings, Movie Revenues, and Viewer Satisfaction. *Journal of Marketing* **74**(1) 108–121.

Mudambi, Susan M. and David Schuff. 2010. What Makes a Helpful Review? A Study of Customer Reviews on Amazon.com. *MIS Quarterly* **34**(1) 185–200.

Ni, Jianmo, Jiacheng Li, and Julian McAuley. 2019. Justifying Recommendations Using Distantly-Labeled Reviews and Fine-Grained Aspects. In *Proceedings of the 2019 Conference on Empirical Methods in*

Park, Aziz, and Lee: *Do Incentivized Reviews Poison the Well?*
Article submitted to *BU Platform Strategy Research Symposium*; manuscript no. 50

19

Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP) 188-197.

Nickerson Raymond S. 1998. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology* **2**(2) 175–220.

Pattabhiramaiah, Adithya, Eric Overby, and Lizhen Xu. 2021. Spillovers from Online Engagement: How a Newspaper Subscriber's Activation of Digital Paywall Access Affects Her Retention and Subscription Revenue. *Management Science* **0**(0).

Perez, Sarah. 2016. Amazon Makes Good on its Promise to Delete "Incentivized" Reviews. Accessed May 23, 2022, https://techcrunch.com/2016/11/23/amazon-made-good-on-its-promise-to-delete-incentivized-reviews/.

Proserpio, Davide, Brett Hollenbeck, and Sherry He. 2020. How Fake Customer Reviews Do — and Don't — Work, *Harvard Business Review*

Puranam, Dinesh, Vrinda Kadiyali, and Vishal Narayan. 2021. The Impact of Increase in Minimum Wages on Consumer Perceptions of Service: A Transformer Model of Online Restaurant Reviews. *Marketing Science* **40**(5) 985-1004.

Qiao, Dandan, Shun-Yang Lee, Andrew B. Whinston, Qiang Wei. 2020. Financial Incentives Dampen Altruism in Online Prosocial Contributions: A Study of Online Reviews. *Information Systems Research* **31**(4) 1361-1375.

Salehan, Mohammad and Dan J. Kim. 2016. Predicting the Performance of Online Consumer Reviews: A Sentiment Mining Approach to Big Data Analytics. *Decision Support Systems* **81** 30–40.

Schindler, Robert M. and Barbara Bickart. 2012. Perceived Helpfulness of Online Consumer Reviews: The Role of Message Content and Style. *Journal of Consumer Behaviour* **11** 234–243.

Wang, Jing, Anindya Ghose, and Panos Ipeirotis. 2012. Bonus, Disclosure, and Choice: What Motivates the Creation of High-Quality Paid Reviews? In *Proceedings of the 33rd International Conference on Information Systems, Orlando, Florida.*

Wang, Shuting Ada, Paul A. Pavlou, and Jing Gong. 2016. On Monetary Incentives, Online Product Reviews, and Sales. In *Proceedings of the 37th International Conference on Information Systems, Dublin, Ireland.*

Wang, Yang, Vandana Ramachandran, and Olivia R. Liu Sheng. 2021. Do Fit Opinions Matter? The Impact of Fit Context on Online Product Returns. *Information Systems Research* **32**(1) 268-289.

Wu, Fang and Bernardo A. Huberman . 2008. How Public Opinion Forms. C. Papadimitrou, S. Zhang, eds. *Internet and Network Economics.* Lecture Notes in Computer Science, Vol. 5385. Springer, Berlin, 334–341.

Xu, Yiqing. 2017. Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models. *Political Analysis* **25**(1) 57–76.

Yin, Dezhi, Sabyasachi Mitra, and Han Zhang. 2016. When Do Consumers Value Positive vs. Negative Reviews? An Empirical Investigation of Confirmation Bias in Online Word of Mouth. *Information Systems Research* **27**(1). 131-144.

Yin, Dezhi, Samuel D. Bond and Han Zhang. 2014. Anxious or Angry? Effects of Discrete Emotions on the Perceived Helpfulness of Online Reviews. *MIS Quarterly* **38**(2) 539-560.

Yin, Dezhi, Samuel D. Bond and Han Zhang. 2021. Anger in Consumer Reviews: Unhelpful but Persuasive. *MIS Quarterly* **45**(3) 1059-1086.

Yu, Yinan, Warut Khern-am-nuai and Alain Pinsonneault. 2022. When Paying for Reviews Pays Off: The Case of Performance-Contingent Monetary Rewards. *MIS Quarterly* **46**(1) 609-626.

Zhu, Feng and Xiaoquan (Michael) Zhang. 2010. Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics. *Journal of Marketing* **74**(2) 133-148.

Zhuang, Mengzhou, Geng Cui, and Ling Peng. 2018. Manufactured Opinions: The Effect of Manipulating Online Product Reviews. *Journal of Business Research* **87** 24-35.