The Impact of Self-awareness of Being Observed on Patient Content Generation: An Empirical Examination from a Quasi Experiment

Mingye Hu
Harbin Institute of Technology
humingyezi@gmail.com

Xiaoxiao Liu
Harbin Institute of Technology
xiaoxiaoliuhit@gmail.com

Bo Xiao
University of Hawaii at Manoa
boxiao@hawaii.edu

Jingbo Shao
Harbin Institute of Technology
boboshao@163.com

Abstract

Though recent research demonstrates the impact of patient generated content on patient outcomes and doctor performance, we still have a limited understanding about how patient content is generated in the first place. In this research, we examine how patients' self-awareness of being observed by their own doctors in online healthcare platform influences patient generated content, including how much they generate and what they generate. Focusing on a leading online healthcare platform, we construct a panel dataset of patient generated content for a matched set of doctors. We find that patients' self-awareness of being observed can increase the quantity of patient generated content. Specially, "being observed" leads to more subjective content, while it has no relationship with objective content. Our results also demonstrate that the mechanism of "being observed" benefits the review quantity at the cost of review quality. We also discuss contributions to user generated content and online healthcare.

1. Introduction

User generated content (UGC) has been considered an important source of information for consumers' purchase decisions and companies' performance across various contexts, including e-commerce, finance markets and stock markets [e.g., 5, 60, 65]. In the healthcare domain, healthcare service is considered as a type of credence goods [21]. That is, doctors know more about patients' conditions and the appropriate treatments, but the patients cannot easily evaluate the appropriateness of the services provided by the doctors [41]. Doctors may utilize the information asymmetry caused by the characteristics of credence goods to provide overtreatment, undertreatment or overcharging [21, 41]. Patient generated content (PGC) (e.g., treatment process, doctor-related information and attitude toward the doctors in the form of review) has been considered as an important factor for patients to discern a doctor's quality and draw numerous researchers to study [22, 25]. Recent studies on PGC have primarily focused on the quality and consequence of PGC, including the relationship between PGC and quality of care [43], and the effects of PGC on patient outcome and doctor performance [41, 67]. However, there exists a limited understanding of how patient content is generated in the first place.

In this paper, we provide insights into the antecedents of patient generated content. Though several studies in other contexts demonstrate that individual characteristics and social influence (e.g., social connection and social ties) can influence user generated content, our work is distinctive in that we examine a possible new driver in the generation of patient content, that is, how patients' self-awareness of being observed by their own doctors on the Internet (i.e. the "being observed effect") influences patient content generation, including how much content the patients' generate and what content they generate. Specifically, we explore how this online "being observed effect" influences the quantity (i.e. volume) and quality (text-based characteristics, e.g., objectivity vs. subjectivity) of patient generated content. Therefore, we seek to answer the following question:

How does patients' self-awareness of being observed by their own doctors in online healthcare platform influences their content generation behavior, including the volume of content they generate and the text-based characteristics of the content they generate?

We examine the "being observed effect" in patient content generation using data obtained from one large online healthcare platform. The challenge to credible causal inference is the endogeneity of patients' self-awareness of being observed. For example, patients' offline experiences with the doctor, which are unobservable to us in this research, may affect patients' self-awareness of being observed and content generation behavior simultaneously. To address this problem, we use a "function launch event" that triggers

URI: https://hdl.handle.net/10125/59522
ISBN: 978-0-9981331-2-6

patients' self-awareness of being observed by their own doctors, and design a quasi-experiment to estimate the causal impact of the "being observed effect" on patient content generation, employing a combination of propensity score matching (PSM) and difference-in-difference (DID) estimation [4, 55].

Our results show that patients' self-awareness of being observed by their own doctors can lead to patients generating more content. Specifically, patients' self-awareness of being observed by their own doctors has a positive and significant effect on the subjective content they generate, while the "being observed effect" has no effect on objective content.

This study makes several contributions to the literature of UGC and online healthcare. First, we contribute to UGC literature by studying the impact of "being observed effect" on the generation of content. While existing studies focus on individual characteristics [32, 71] and social influence [24, 62] as antecedent of user content generation, we explore a new mechanism, the "being observed effect," referred specifically to being observed by those who are being reviewed. Second, we extend the research on UGC to the domain of healthcare, especially from the perspective of antecedents. Extant literature in PGC in healthcare focuses on the quality and consequences of PGC; we take an additional step to studying the antecedent of PGC in the healthcare domain.

2. Literature review

We draw on the literature on user generated content (UGC), which has been extensively examined in the forms of online reviews, online rating and word-of-mouth (WOM), in multiple contexts. Our review of the literature reveals three categories of the work in this domain. Literature in the first category seeks to understand the quality and helpfulness of UGC, including whether the UGC is credible and what makes UGC helpful [47, 49, 69]. For example, Mudambi and Schuff [49] explore what factors make reviews helpful. They find that review depth has a positive effect on the helpfulness of the review, while product type (i.e., experience goods vs. search goods) has a moderating impact on the effect of review depth and review extremity on review helpfulness. In the healthcare domain, researchers have explored the relationship between doctors' online rating and offline quality to explore whether the online rating can reflect the true quality of doctors [22, 25, 43]. For example, Lu and Rui [43] study whether online rating can index doctors' medical quality. Using data from RateMDs and hospitals, they find that online doctors' ratings can provide valuable information for patients to judge doctors' medical quality.

In the second category, researchers have examined the impact of UGC, showing that UGC has significant impact on a variety of outcomes, including individual behaviors [29, 50, 54], market performance [1, 17, 33, 35, 60, 65, 72] and social network outcomes [61]. For example, Park, Lee and Han [50] find that the quantity of online consumer reviews has a positive effect on consumers' purchase intention. Trusov, Bucklin and Pauwels [61] study the effect of WOM on member growth in social networks and find that WOM has a strong positive effect on new customer acquisition in the social network. In addition to the quantity of UGC, prior research also explores the effects of different metrics of online consumer reviews on performance across different platforms [see 5 for a review]. The effects of text-based characteristics in UGC (e.g., objective UGC and subjective UGC) have also been explored [9, 16, 23, 31, 36, 40, 59]. For example, Liu, Ozanne and Mattila [40] explore the effectiveness of subjectivity and objectivity expression in online reviews, and find that subjective contents in online reviews can increase men's purchase intention in the hedonic context and women's purchase intention in the utilitarian context. In healthcare domain, existing studies have explored how patient generated content affects patients' outcomes [67] and doctors' performance [41]. For example, Yan et al. [67] study how other patients' comments influence patients' perceived treatment outcome and find that comments with positive sentiment from other patients have a negative effect on the patients' perceived treatment outcome.

The third category, where our own interest primarily lies, is a small but growing body of research that looks at the antecedent of UGC, i.e. what factors affect user content generation behavior. Existing research examines the antecedent from individual and product factors as well as social factors. In terms of individual and product factors, literature shows that individual characteristics, such as gender [71], cultural background [30], experience [53, 68], uniqueness [14], self-needs [2, 64], self-expression [56, 57] and customer type [3, 32, 39, 52] can affect content generation behavior, including volume and text-based characteristics (e.g., positive and negative content). For example, Zhang, Feick and Mittal [71] explore the different impact of gender in negative WOM transmission, and show that the difference is driven by men's concern for self and women's concern for others. In addition, users' content generation behavior can also be influenced by different product types [8, 19], brand [42], content acquisition method [15] and communication channel [7]. For example, Lovett, Peres and Shachar [42] study brand characteristics as antecedent of WOM. They find that brand characteristics, including social, emotional, and functional aspects, have a significant effect on online and offline WOM mentions. In terms of social factors, most of the literature focuses on the social influence effect, i.e. how other behaviors or other audiences in social environments influence user content generation behavior. Existing research shows that user content generation behavior can be influenced by prior UGC [38, 45, 46, 58], audience size [6], social management [44, 63] and online interaction (or social ties) [24, 62, 70]. For example, Lee, Hosanagar and Tan [38] explore the different effects of prior ratings from friends and strangers, and find that higher prior ratings can increase the intention of users to give a higher rating and this effect is weaker when the prior ratings are from friends. Goes, Lin and Au Yeung [24] study the impact of online interaction on user content generation behaviors. Using data from a product review website, they find that when the users become more popular (i.e., more followers), they generate more reviews and more objective reviews. A few research studies focus on the antecedent of UGC from the perspective of anonymity and social presence, i.e. how personal social exposure affects user content generation behavior. Huang, Hong and Burtch [34] explore the effect of social presence on users' content generation by studying the social network integration in Yelp.com and TripAdvisor.com. They find that by increasing social presence, social network integration can lead to more UGC volume and more emotional UGC, while decreasing cognitive language, negative emotion and expression of disagreement words.

Our research falls under the third category, and we seek to fill two critical gaps in literature. First, even though prior literature advances our understanding of the UGC generation behavior, it is mostly restricted to dominant contexts such as e-commerce, films, restaurants, stock markets and finance markets. There is limited research in the context of credence goods (e.g., healthcare service) to explore factors driving patients to generate online content. Healthcare service is a typical credence good in that, while doctors know about a patient's condition and appropriate treatment, the patients cannot evaluate the appropriateness of the services provided by the doctors [21]. Therefore, other patients' content on the Internet is an extremely important information source for patients to discern doctors' quality. Given this background, examining the motivation of patient content generation is important in the healthcare domain.

Second, our study examines the effect of the reviewers' (i.e. patients) self-awareness of being observed by the specific people who are being reviewed (i.e. doctors), a potential new driver of online content generation that has not previously been identified. That is, how do patients change their content generation behavior when they feel they are being observed by their own doctors? This is a unique mechanism which is similar to but not equivalent to non-anonymous. It only increases the patients' feeling that they are becoming being observed by their own doctors, because their doctors can track their generated content and may know who they are. This lead patients to be non-anonymous to the specific group in the platform (i.e., their own doctors). However, the patients are still being kept anonymous to other users, including all the patients and other doctors. This is different from the mechanism examined in Huang, Hong and Burtch [34] in which the users are non-anonymous to all users. This unique setup can allow us to examine patients' self-awareness of being observed by their own doctors on their content generation behaviors, including how much they produce (quantity of PGC) and what they produce (objective content vs subjective content).

3. Hypotheses development

3.1. The quantity of PGC

The quantity of PGC reflects a doctor's popularity, since it is reasonable to assume that the quantity of PGC is related to the number of patients who have chosen this doctor. Patients' self-awareness of being observed by their doctors, through enhancing patients' sense of presence, may affect their decisions to contribute contents.

First, patients' self-awareness of being observed by their doctors enhances their sense of presence as unique individuals to their doctors, increasing their feelings of connection to the doctor they are reviewing [13]. Patients know that the doctor could trace back from the content and obtain their personal information (e.g. real name, cell phone, and even treatment records). As such, patients are more likely to participate actively in online healthcare platform to get the doctors' attention and hopefully strengthen their connection with the doctors, which is beneficial to their own treatment process. Patients may also believe that online and offline interaction would provide their doctors with more opportunities to know them, a belief that may also encourage them to generate more content.

Second, patients' self-awareness of being observed by their doctors motivates them to act prosocially to gain a good impression in the eyes of observers [51]. Using online healthcare platforms, patients could receive or give social support, including informational support in the form of sharing advice or referrals, and emotional support in the form of sharing happiness or sadness [66]. When observed by their doctors, patients would be more willing to help peer patients by providing informational and emotional support, resulting in their contributing larger volume of content. Further, patients are also more likely to write PGC for giving feedback to their doctors, aiming to encourage them or help to improve their service, which will be beneficial to the relationship between patients and doctors in the long run [48]. Even for patients receiving poor services and treatments, switching to other doctors would require extra cost of time and energy, which some of them may not want to load, especially for patients with a limited choice of doctors. As such, unsatisfied patients may use this new channel to communicate with their doctors in our context, instead of keeping silent. Hence, most of patients tend to increase their content generation behaviors when being observed by their doctors. Accordingly, we propose the following hypothesis:

H1: Patients' self-awareness of being observed by their doctors increase the quantity of patient generated content.

3.2. The quality of PGC

In this study, the quality of PGC is defined as the information quality of PGC from the perspective of text-based characteristics (i.e. subjective and objective). Studies in marketing show that objective reviews are more effective than subjective ones, since the former contains more specific and clearer opinions [50]. In healthcare setting, we consider objective PGC high-quality PGC, as it is based on specific facts about the process of healthcare services. In contrast, subjective PGC is considered low-quality PGC, which is based on emotion as opposed to reasoned arguments. We argue that patients' self-awareness of being observed by their doctors affect the quality of their contents.

When patients are aware of being observed by their doctors, their sense of presence in front of their doctors is getting higher. Sense of presence in social contexts influences the extent to which one displays emotions. Situations in which others are present or only imaginary present affect the amount of emotion expression [20]. For example, Huang, Hong and Burtch [34] have shown that social presence of friends in online platforms increases language reflecting affective processes in review text, compared with cognitive processes. Affective processes include one's feelings related to the object of being evaluated. In the same line of reasoning, patients' self-awareness of being observed of their doctors motivate them displaying more emotional expressions in PGC, through either a positive or a negative tone, by which increase the subjectivity of PGC and decline the objectivity of PGC relatively. Therefore, we propose the following two hypothesis:

H2a: Patients' self-awareness of being observed by their doctors leads to more subjective patient generated content.

H2b: Patients' self-awareness of being observed by their doctors leads to less objective patient generated content.

4. Research setting

4.1. Research context

We collected data from a leading Chinese online healthcare platform, which displays information about doctors from a variety of hospitals across China. An information page is created by the platform for each listed on the platform. On this page, visitors can see detailed information about the doctor (e.g., departments, title, specialty and outpatient schedule) and his/her affiliated hospitals (e.g., telephone, rank and address). Patients can generate content (e.g., treatment process, doctor-related information and attitude toward the doctors) in the form of review about the doctors they have seen before. Before they generate and publish the content, they must register with the platform. However, the platform partially masks the user's ID in the published content. Moreover, the platform only allows patients to register and log in, and does not provide any channels for doctors to register and participate in the online platform.

In order for doctors to participate in and utilize online platform to manage their patients and learn knowledge, in March 2008, this online healthcare platform implemented a new feature that allows doctors to create their homepages to register and log in. Doctors can update their personal information and outpatient schedule in their homepage. The creation of homepages allows doctors, when logged in, to track and check their patients' generated content instantly. For the patients, they generate contents in the doctors' information pages if the doctors do not create homepages. After the doctors create their homepages, there is a button link (homepage) in the doctors' information pages for patients to distinguish and identify these doctors. Therefore, patients can easily know whether their doctors have created homepages and logged in. Thus, doctors' creation of homepage may increase their patients' feeling that they are becoming non-anonymous to their own doctors (being observed). Patients, however, can generate content to evaluate their doctors in the information page whether the doctors create homepage or not.

4.2. Identification strategy

To establish a causal relationship between patients' self-awareness of being observed and their content generation behavior, we utilize the launch of the function, i.e. "creation of homepage", to build a quasi-experimental research design. That is, the creation of homepage makes patients know whether their doctors have logged in the platform and tracked their generated content, which may increase the feeling of being observed by their own doctors. For doctors who created their homepage, the creation of homepages would increase their patients' self-awareness of being observed. The patients of doctors who did not create their own homepage would not be affected by the launch of this function. Therefore, we have two distinct groups of doctors, where a "treatment" group contains doctors who created their homepage and a "control" group that contains doctors who did not create the homepage. In order to mimic a random experimental design and get an unbiased estimate of "treatment effect", we utilized PSM and DID estimation [18]. By using these methods, we hope to solve the endogeneity issues by controlling for self-selection.

4.3. Data collection

We used a web crawler to collect data on two diseases: fracture and coronary heart disease, from September 2007 to August 2008. The data includes the doctor's title, the rank of the hospital with which the doctor is affiliated, the doctor's geographic location, a record of patient generated content, and the date of homepage creation. We obtained a sample of 2055 doctors with 297 doctors in the treatment group and 1758 doctors in the control group. Different doctors in the treatment group created their homepages at different points of time during the study period.

4.4. Variable operationalization

4.4.1. Dependent variables. We explore patient content generation behavior from two aspects: patient generated content quantity and patient generated content quality. PGC quantity is the volume of PGC and is denoted as PGC_Volume_{jt}, which is calculated as the total number of patient generated content data points with doctor *j* in the period *t*. Prior literature has shown that high quality or useful UGC usually includes objective information that is less emotionally expressive [16, 30, 34, 50, 69]. Therefore, we use text-based features, i.e. objectivity and subjectivity, as the criterion to assess PGC quality. Objective PGC is content that mainly contains objective information, is understandable, and most importantly, has detailed information about treatment process and doctors. Subjective PGC is content that mainly contains emotional, subjective information, and has no detailed information about treatment processes and doctors. We use artificial classification and machine learning to classify the data into objectivity and subjectivity. For example, "Bilateral knee joint replacement surgery. This surgery took about 2 hours, a small wound. My blood loss was below 200ml and now postoperative recovery is good." is objective content, while "The doctor is great. He is professional, easygoing and patient. My benefactor!" is subjective content. Therefore, we have two variables to measure PGC quality: objective PGC, which is denoted as $PGC\_Objectivity_{jt}$ and calculated as the number of objective patient generated content data points with doctor *j* in the period *t*; subjective PGC, which is denoted as $PGC\_Subjectivity_{jt}$ and calculated as the number of subjective patient generated content with doctor *j* in the period *t*.

4.4.2. Independent variables. We created a binary variable $DParT_{jt}$ to capture the periods before and after the doctor *j* creates a homepage. The variable is one if the period is after the doctor *j* creates homepage at the given time *t*. It is zero if the period is before the doctor *j* creates homepage at the given time *t*. We also created a treatment dummy $TreatD_j$ to capture if a doctor is in the treatment or control group. The variable is one if the doctor is in the treatment group. It is zero if the doctor is in the control group.

4.4.3. Control variables. We also included several control variables in our model, including the doctor's title ($DTitle\_D_{1j}$ takes the value one for "chief doctor", $DTitle\_D_{2j}$ takes the value one for "associate chief doctor", and zero for other doctors), the doctor's hospital rank (denoted as $HLevel_j$, takes the value one for the highest ranked hospitals, and zero for lower ranked hospitals), the GDP of the city where the doctor *j* is located (denoted as GDP_j), and disease type (denoted as $Disease_j$, takes the value one for coronary heart disease, and zero for fracture disease). These control variables were entered in the PSM.

5. Data analysis

We have two groups (i.e. treatment group and control group) according to the identification strategy. DID analysis calculated the effect of treatment by comparing the outcome of the treatment and control groups in the pre- and post- treatment (i.e. the creation of homepage), which helps us mitigate the effects of extraneous factors [4, 55]. We used PSM to select a group of doctors in the control group who are comparable to the doctors in the treatment group in terms of the doctors' background variables, so that the differences in the outcome variables cannot be attributed to the differences in doctors' background. We then ran a DID model to test the causal impact of patients' self-awareness of being observed on the dependent variables.

5.1. Propensity score matching

We conducted propensity score matching following the standard steps outlined in the prior literature [11, 28, 55]. First, we used a logistic model that includes the doctors' background variables to estimate the propensity scores (see Table 1). Second, we matched the doctors in the treatment and control groups using the nearest neighborhood without caliper pair matching algorithm. The PSM generated 292 doctors in the treatment group and 292 doctors in the control group. Third, we checked if the common support requirement is met by plotting the propensity score distributions through histogram plots and box plots [28, 37, 55] (see Figure 1). It can be seen that the propensity score distributions for treatment and control groups are different before matching. However, after matching, the propensity score distributions for the treatment and control groups are almost identical. Therefore, we were confident that the matching results met the common support requirement, and concluded that the treatment and control groups have no significant difference in the propensity score. Fourth, we checked the matching quality to see if the two groups are balanced on the covariates by comparing the covariates between treatment and control groups before and after matching (see Table 2). The results show that, after matching, the treatment and control groups have no significant differences on the covariates.

Table 1. Logistic regression model

Variable	Coefficient	Std. error
DTitle_D1	-1.236***	0.259
DTitle_D2	-0.805**	0.267
HLevel	-0.555**	0.203
GDP	-0.181**	0.064
Disease	-0.417***	0.131
Constant	1.529**	0.566
Log likehood	-813.701	

Note: *** p<0.001; ** p<0.01; * p<0.05

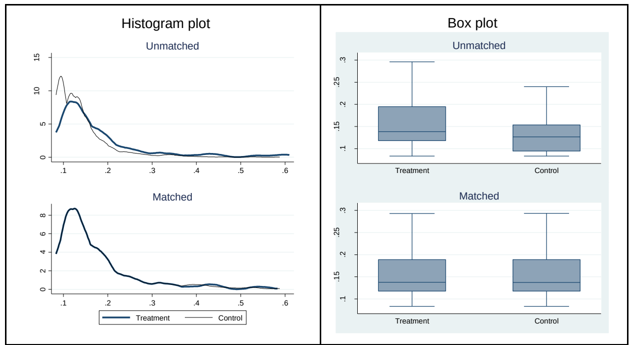

Figure 1. Distribution of propensity score before and after matching

Table 2. Covariate comparison before and after matching

		Mean		T-value
Variable		Treatment	Control	
DTitle_D1	Unmatched	0.586	0.739	-5.47
	Matched	0.596	0.596	-0.00
DTitle_D2	Unmatched	0.317	0.234	3.04
	Matched	0.322	0.322	0.00
HLevel	Unmatched	0.862	0.938	-4.68
	Matched	0.877	0.894	-0.65
GDP	Unmatched	8.252	8.503	-4.30
	Matched	8.287	8.335	-0.57
Disease	Unmatched	0.508	0.639	-4.30
	Matched	0.517	0.483	0.83

5.2. Difference-in-difference analysis

The DID models of patients content generation behavior are specified as follows.

PGC quantity model (1):

$$PGC\_Volume_{ijt} = \alpha_j + \beta_1 DParT_{ijt} + \beta_2 TreatD_j \times DParT_{ijt} + \varepsilon_{jt}$$

PGC quality model (2):

$$PGC\_Quality_{ijt} = \alpha_j + \beta_1 DParT_{ijt} + \beta_2 TreatD_j \times DParT_{ijt} + \varepsilon_{jt}$$

where i denotes a matched pair of doctors, j denotes a treatment or control group doctor, and t denotes the time period. $TreatD_j$ is the treatment dummy that indicates whether doctor j is in the treatment group ($TreatD_j$=1) or the control group ($TreatD_i$=0). $DParT_{ijt}$ is a dummy variable that indicates if the period is before ($DAppT_{ijt}$=0) or after the launch of the mobile app ($DAppT_{ijt}$=1), respectively, for doctors belonging to the matched pair i. α_j is the doctor fixed effects that help to control for the unobserved heterogeneity across doctors. $PGC\_Quality_{ijt}$ are the text-based characteristics including PGC objectivity and PGC subjectivity.

6. Results

Table 3 shows the results of DID estimation. The parameter corresponding to the treatment effect of being observed is positive and significant in the PGC quantity model (coefficient=0.107, p-value=0.000, Model 1). This suggests that patients' self-awareness of being observed by their own doctors has a positive effect on their generated content quantity. Thus, we find support for H1. For the PGC quality model, patients' self-awareness of being observed by their own doctors has a positive and significant effect on the PGC subjectivity (coefficient=0.100, p-value=0.000, Model 3), which suggests that patients would generate more subjective content when they feel they are being observed by their doctors. Thus, our H2a are supported. However, patients' self-awareness of being observed by their own doctors has no significant effect on the PGC objectivity (coefficient=0.012, p-value=0.081, Model 2). This suggests that when patients feel they have been observed by their own doctors, they do not change their behavior to post more objective content. Thus, H2b is not supported.

Table 3. Results

	Quantity	Quality	
variables	*Volume*	*Objectivity*	*Subjectivity*
Models	(1)	(2)	(3)
DParT	-0.042*** (0.009)	-0.004 (0.004)	-0.037*** (0.008)
TreatD×*DParT*	0.107*** (0.015)	0.012 (0.007)	0.100*** (0.014)
Constant	0.104*** (0.003)	0.021*** (0.001)	0.086*** (0.002)
Doctor fixed effects	Y	Y	Y
Clustered Errors	Y	Y	Y
Number of doctors	584	584	584
R-squared	0.230	0.144	0.207

Note: ***p<0.001, **p<0.01 and *p<0.05.

7. Robustness checks

7.1. Matching with alternative techniques

We employed other matching algorithms to verify the robustness of our results, that is, optimal pair matching and nearest neighborhood with caliper (0.25*SD, SD is the standard deviation of propensity score) pair matching. We find that the estimations are largely consistent with our main results.

7.2. Robustness of the DID analysis

We first used a relative time model [10, 26, 27] to check the parallel trend assumption of DID estimation, which requires that there is no pre-treatment heterogeneity in the trends between treatment and control groups [4]. Specifically, we created a series of time dummies to indicate the relative chronological distance between the period *t* and the treatment time (i.e. the launch of homepages creation), following prior literature [4, 10, 12, 26, 27, 34]. This approach can help determine the existence of pre-treatment heterogeneity in the trends between treatment and control groups (i.e. a significant difference between treatment and control groups before the treatment). Therefore, we specified the following models:

$$PGC\_Volume_{ijt} = \alpha_j + \beta_1 Time\_Dummies_t + \beta_2 TreatD_j \times Time\_Dummies_t + \varepsilon_{jt} \quad (3)$$

$$PGC\_Quality_{ijt} = \alpha_j + \beta_1 Time\_Dummies_t + \beta_2 TreatD_j \times Time\_Dummies_t + \varepsilon_{jt} \quad (4)$$

We drew the coefficients (β_2) of each $TreaD_j \times Time\_Dummies_t$ for our dependent variables from the above estimation in Figure 2. As shown in the figure, there is no evidence of significant pre-treatment difference in the pre-treatment periods, which supports the parallel trend assumption.

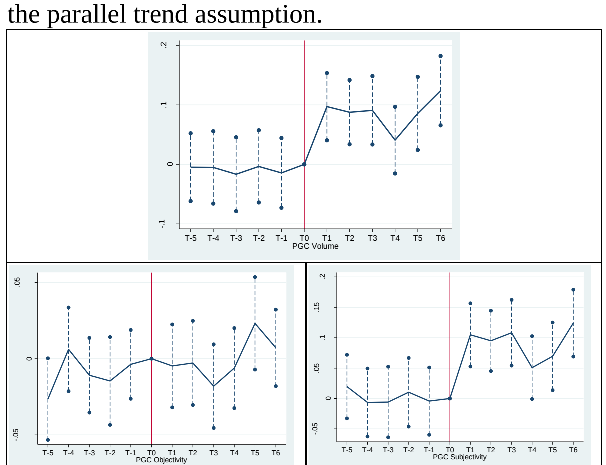

Figure 2. The coefficients of treatment effects in different periods

Second, in our main analysis, we used data that contains six months before and after the launch of the function (creation of homepages). For the robustness test, we ran the DID analysis using different periods before and after the launch of the function (i.e. data contains five, three months before and after, respectively) to confirm that our results are not caused by unobservable factors in certain periods and make sure the results are robust to the different time windows [55]. We found that the results using different time windows are similar to the main results.

8. Discussion

This research examines the antecedent of patient generated content in an online healthcare platform to explore how the "being observed effect" influences the volume and types of the contents that patients generate. By building a quasi-experimental design, we find that patients' self-awareness of being observed by their own doctors can cause them to generate more reviews about their doctors. This shows that the mechanism of "being observed" can benefit doctors and the platform by increasing the quantity of patient generated content, such as online reviews. Specifically, the results also show that the "being observed effect" can stimulate patients to generate more subjective content. However, this mechanism has no relationship with objective content generation. A possible explanation for this unexpected finding is that "being observed" may have motivated patients to generate content as a way to communicate and build the relationship with their doctors, instead of using it as a traditional UGC to help other patients. As increased subjectivity indicates low quality of patient generated content, the mechanism of "being observed" may turn out to be harmful to the platform by increasing the proportion of low quality patient generated content.

8.1. Theoretical contribution

First, the current project joins the small but growing literature that examines the antecedent of user generated content. Exploring the factors driving user content generation has been a prominent research area in the field of Information Systems. Existing literature has primarily focused on individual and product factors, such as gender, culture background and product types [32, 71], as well as social influence factors, such as prior UGC, audience size and social connection [24, 62]. We contribute to UGC literature by examining a new driver, i.e. the "being observed effect" (in the sense of being observed by the person the user is generating content about, or the people being reviewed.). Our empirical study has established a causal link between "being observed effect" and generation of content.

Second, existing literature in PGC in healthcare focuses on its quality and consequence. We extend this literature by studying the generation of PGC in the healthcare domain, especially from the perspective of antecedent factors.

8.2. Practical contribution

Our results have a number of implications for practice. First, our results can provide important insights into the cultivation and accumulation of patient generated content in online healthcare platforms. We show that patients' "being observed effect" can incentivize patients to generate more content. This indicates that the mechanism of "being observed by the people who are reviewed" is a useful tool to increase the volume of PGC, which is an important resource for review or rating websites. Second, we show that the mechanism benefits the review quantity at the cost of review quality (when quality is indicated by content objectivity). Therefore, online healthcare platforms should consider designing additional communication functions to encourage patients to contribute more objective content, thus increasing the quality of patient generated content.

8.3. Limitations

Our study has several limitations. First, we only analyzed two diseases. Future studies may consider different types of medical conditions and compare the effects of "being observed effect" on these different conditions. Second, the text-based characteristics in this study only look at objectivity and subjectivity. Patient content usually contains various types of information, such as treatment outcomes, prior treatment experience and information about hospitals. Further studies can classify the content into more types to study how "being observed effect" affects detailed information in patient generated content.

9. Acknowledgments

This study was funded by the National Natural Science Foundation of China Grants (71672048,71172155).

10. References

[1] Aggarwal, R., R. Gopal, A. Gupta, and H. Singh, *Putting money where the mouths are: The relation between venture financing and electronic word-of-mouth.* Information Systems Research, 2012. 23(3-part-2): p. 976-992.
[2] Alexandrov, A., B. Lilly, and E. Babakus, *The effects of social-and self-motives on the intentions to share positive and negative word of mouth.* Journal of the Academy of Marketing Science, 2013. 41(5): p. 531-546.
[3] Anderson, E.T. and D.I. Simester, *Reviews without a purchase: Low ratings, loyal customers, and deception.* Journal of Marketing Research, 2014. 51(3): p. 249-269.
[4] Angrist, J.D. and J.-S. Pischke, *Mostly harmless econometrics: An empiricist's companion*. 2008, Princeton, New Jersey: Princeton university press.

[5] Babić Rosario, A., F. Sotgiu, K. De Valck, and T.H. Bijmolt, *The effect of electronic word of mouth on sales: A meta-analytic review of platform, product, and metric factors.* Journal of Marketing Research, 2016. 53(3): p. 297-318.
[6] Barasch, A. and J. Berger, *Broadcasting and narrowcasting: How audience size affects what people share.* Journal of Marketing Research, 2014. 51(3): p. 286-299.
[7] Berger, J. and R. Iyengar, *Communication channels and word of mouth: How the medium shapes the message.* Journal of Consumer Research, 2013. 40(3): p. 567-579.
[8] Berger, J. and E.M. Schwartz, *What drives immediate and ongoing word of mouth?* Journal of Marketing Research, 2011. 48(5): p. 869-880.
[9] Berger, J., A.T. Sorensen, and S.J. Rasmussen, *Positive effects of negative publicity: When negative reviews increase sales.* Marketing Science, 2010. 29(5): p. 815-827.
[10] Burtch, G., S. Carnahan, and B.N. Greenwood, *Can you gig it? An empirical examination of the gig economy and entrepreneurial activity.* Management Science, 2018. https://doi.org/10.1287/mnsc.2017.2916.
[11] Caliendo, M. and S. Kopeinig, *Some practical guidance for the implementation of propensity score matching.* Journal of Economic Surveys, 2008. 22(1): p. 31-72.
[12] Chan, J. and A. Ghose, *Internet's dirty secret: assessing the impact of online intermediaries on HIV transmission.* MIS Quarterly, 2014. 38(4): p. 955-976.
[13] Chatting, D.J., J.S. Galpin, and J.S. Donath. *Presence and portrayal: video for casual home dialogues*. in *Proceedings of the 14th ACM international conference on Multimedia*. 2006. ACM.
[14] Cheema, A. and A.M. Kaikati, *The effect of need for uniqueness on word of mouth.* Journal of Marketing Research, 2010. 47(3): p. 553-563.
[15] Chen, Z. and J. Berger, *How content acquisition method affects word of mouth.* Journal of Consumer Research, 2016. 43(1): p. 86-102.
[16] Chen, Z. and N.H. Lurie, *Temporal contiguity and negativity bias in the impact of online word of mouth.* Journal of Marketing Research, 2013. 50(4): p. 463-476.
[17] Chevalier, J.A. and D. Mayzlin, *The effect of word of mouth on sales: Online book reviews.* Journal of Marketing Research, 2006. 43(3): p. 345-354.
[18] Dehejia, R.H. and S. Wahba, *Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs.* Journal of the American Statistical Association, 1999. 94(448): p. 1053-1062.
[19] Dellarocas, C., G. Gao, and R. Narayan, *Are consumers more likely to contribute online reviews for hit or niche products?* Journal of Management Information Systems, 2010. 27(2): p. 127-158.
[20] Derks, D., A.H. Fischer, and A.E. Bos, *The role of emotion in computer-mediated communication: A review.* Computers in Human Behavior, 2008. 24(3): p. 766-785.
[21] Dulleck, U. and R. Kerschbamer, *On doctors, mechanics, and computer specialists: The economics of credence goods.* Journal of Economic Literature, 2006. 44(1): p. 5-42.
[22] Gao, G.G., B.N. Greenwood, R. Agarwal, and J.S. McCullough, *Vocal Minority and Silent Majority: How Do Online Ratings Reflect Population Perceptions of Quality.* MIS Quarterly, 2015. 39(3): p. 565-589.
[23] Ghose, A. and P.G. Ipeirotis, *Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics.* IEEE Transactions on Knowledge and Data Engineering, 2011. 23(10): p. 1498-1512.
[24] Goes, P.B., M. Lin, and C.-m. Au Yeung, *"Popularity effect" in user-generated content: Evidence from online product reviews.* Information Systems Research, 2014. 25(2): p. 222-238.
[25] Gray, B.M., J.L. Vandergrift, G.G. Gao, J.S. McCullough, and R.S. Lipner, *Website ratings of physicians and their quality of care.* JAMA Internal Medicine, 2015. 175(2): p. 291-293.
[26] Greenwood, B.N. and R. Agarwal, *Matching platforms and HIV incidence: An empirical investigation of race, gender, and socioeconomic status.* Management Science, 2015. 62(8): p. 2281-2303.
[27] Greenwood, B.N. and S. Wattal, *Show Me the Way to Go Home: An Empirical Investigation of Ride-Sharing and Alcohol Related Motor Vehicle Fatalities.* MIS Quarterly, 2017. 41(1): p. 163-187.
[28] Guo, S. and M.W. Fraser, *Propensity Score Analysis: Statistical Methods and Applications*. Vol. 11. 2014: SAGE Publications.
[29] Hennig-Thurau, T., C. Wiertz, and F. Feldhaus, *Does Twitter matter? The impact of microblogging word of mouth on consumers' adoption of new movies.* Journal of the Academy of Marketing Science, 2015. 43(3): p. 375-394.
[30] Hong, Y., N. Huang, G. Burtch, and C. Li, *Culture, conformity, and emotional suppression in online reviews.* Journal of the Association of Information Systems, 2016. 17(11): p. 737-758.
[31] Hu, N., N.S. Koh, and S.K. Reddy, *Ratings lead you to the product, reviews help you clinch it? The mediating role of online review sentiments on product sales.* Decision Support Systems, 2014. 57: p. 42-53.
[32] Hu, N., P.A. Pavlou, and J. Zhang, *On self-selection biases in online product reviews.* MIS Quarterly, 2017. 41(2): p. 449-471.
[33] Huang, L., J. Zhang, H. Liu, and L. Liang, *The effect of online and offline word-of-mouth on new product diffusion.* Journal of Strategic Marketing, 2014. 22(2): p. 177-189.
[34] Huang, N., Y. Hong, and G. Burtch, *Social Network Integration and User Content Generation: Evidence from Natural Experiments.* MIS Quarterly, 2017. 41(4): p. 1035-1058.
[35] Kwark, Y., J. Chen, and S. Raghunathan, *Online product reviews: Implications for retailers and competing manufacturers.* Information Systems Research, 2014. 25(1): p. 93-110.
[36] Lappas, T., G. Sabnis, and G. Valkanas, *The impact of fake reviews on online visibility: A vulnerability assessment of the hotel industry.* Information Systems Research, 2016. 27(4): p. 940-961.
[37] Lechner, M., *Some practical issues in the evaluation of heterogeneous labour market programmes by matching methods.* Journal of the Royal Statistical Society: Series A (Statistics in Society), 2002. 165(1): p. 59-82.
[38] Lee, Y.-J., K. Hosanagar, and Y. Tan, *Do I follow my friends or the crowd? Information cascades in online movie ratings.* Management Science, 2015. 61(9): p. 2241-2258.

[39] Li, X. and L.M. Hitt, *Self-selection and information role of online product reviews.* Information Systems Research, 2008. 19(4): p. 456-474.
[40] Liu, S.Q., M. Ozanne, and A.S. Mattila, *Does expressing subjectivity in online reviews enhance persuasion?* Journal of Consumer Marketing, 2018. https://doi.org/10.1108/JCM-02-2017-2109.
[41] Liu, X., X. Guo, H. Wu, and T. Wu, *The Impact of Individual and Organizational Reputation on Physicians' Appointments Online.* International Journal of Electronic Commerce, 2016. 20(4): p. 551-577.
[42] Lovett, M.J., R. Peres, and R. Shachar, *On brands and word of mouth.* Journal of Marketing Research, 2013. 50(4): p. 427-444.
[43] Lu, S.F. and H. Rui, *Can We Trust Online Physician Ratings? Evidence from Cardiac Surgeons in Florida.* Management Science, 2017. 64(6): p. 2557 - 2573.
[44] Miller, A.R. and C. Tucker, *Active social media management: the case of health care.* Information Systems Research, 2013. 24(1): p. 52-70.
[45] Moe, W.W. and D.A. Schweidel, *Online product opinions: Incidence, evaluation, and evolution.* Marketing Science, 2012. 31(3): p. 372-386.
[46] Moe, W.W. and M. Trusov, *The value of social dynamics in online product ratings forums.* Journal of Marketing Research, 2011. 48(3): p. 444-456.
[47] Moore, S.G., *Attitude predictability and helpfulness in online reviews: the role of explained actions and reactions.* Journal of Consumer Research, 2015. 42(1): p. 30-44.
[48] Moyer, C.A., D.T. Stern, K.S. Dobias, D.T. Cox, and S.J. Katz, *Bridging the electronic divide: patient and provider perspectives on e-mail communication in primary care.* The American journal of managed care, 2002. 8(5): p. 427-433.
[49] Mudambi, S.M. and D. Schuff, *Research note: What makes a helpful online review? A study of customer reviews on Amazon. com.* MIS Quarterly, 2010. 34(1): p. 185-200.
[50] Park, D.-H., J. Lee, and I. Han, *The effect of on-line consumer reviews on consumer purchasing intention: The moderating role of involvement.* International Journal of Electronic Commerce, 2007. 11(4): p. 125-148.
[51] Pfattheicher, S. and J. Keller, *The watching eyes phenomenon: The role of a sense of being seen and public self‐awareness.* European Journal of Social Psychology, 2015. 45(5): p. 560-566.
[52] Punj, G.N., *Do consumers who conduct online research also post online reviews? A model of the relationship between online research and review posting behavior.* Marketing Letters, 2013. 24(1): p. 97-108.
[53] Rice, S.C., *Reputation and uncertainty in online markets: an experimental study.* Information Systems Research, 2012. 23(2): p. 436-452.
[54] Riegner, C., *Word of mouth on the web: The impact of Web 2.0 on consumer purchase decisions.* Journal of Advertising Research, 2007. 47(4): p. 436-447.
[55] Rishika, R., A. Kumar, R. Janakiraman, and R. Bezawada, *The effect of customers' social media participation on customer visit frequency and profitability: an empirical investigation.* Information Systems Research, 2013. 24(1): p. 108-127.
[56] Saenger, C., V.L. Thomas, and J.W. Johnson, *Consumption‐focused self‐expression word of mouth: A new scale and its role in consumer research.* Psychology & Marketing, 2013. 30(11): p. 959-970.
[57] Sicilia, M., E. Delgado‐Ballester, and M. Palazon, *The need to belong and self‐disclosure in positive word‐of‐mouth behaviours: The moderating effect of self–brand connection.* Journal of Consumer Behaviour, 2016. 15(1): p. 60-71.
[58] Sridhar, S. and R. Srinivasan, *Social influence effects in online product ratings.* Journal of Marketing, 2012. 76(5): p. 70-88.
[59] Tang, T., E. Fang, and F. Wang, *Is neutral really neutral? The effects of neutral user-generated content on product sales.* Journal of Marketing, 2014. 78(4): p. 41-58.
[60] Tirunillai, S. and G.J. Tellis, *Does chatter really matter? Dynamics of user-generated content and stock performance.* Marketing Science, 2012. 31(2): p. 198-215.
[61] Trusov, M., R.E. Bucklin, and K. Pauwels, *Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site.* Journal of Marketing, 2009. 73(5): p. 90-102.
[62] Wang, C., X. Zhang, and I.-H. Hann, *Socially nudged: A quasi-experimental study of friends' social influence in online product ratings.* Information Systems Research, 2018. https://doi.org/10.1287/isre.2017.0741.
[63] Wang, Y. and A. Chaudhry, *When and How Managers' Responses to Online Reviews Affect Subsequent Reviews.* Journal of Marketing Research, 2018. 55(2): p. 163-177.
[64] Wien, A.H. and S.O. Olsen, *Understanding the relationship between individualism and word of mouth: A self‐enhancement explanation.* Psychology & Marketing, 2014. 31(6): p. 416-425.
[65] Wu, C., H. Che, T.Y. Chan, and X. Lu, *The economic value of online reviews.* Marketing Science, 2015. 34(5): p. 739-754.
[66] Yan, L. and Y. Tan, *Feeling blue? go online: an empirical study of social support among patients.* Information Systems Research, 2014. 25(4): p. 690-709.
[67] Yan, L., X. Yan, Y. Tan, and S.X. Sun, *Shared Minds: How Patients Use Collaborative Information Sharing via Social Media Platforms.* Production and Operations Management, 2018. https://doi.org/10.1111/poms.12895.
[68] Yang, S., M. Hu, R.S. Winer, H. Assael, and X. Chen, *An empirical study of word-of-mouth generation and consumption.* Marketing Science, 2012. 31(6): p. 952-963.
[69] Yin, D., S.D. Bond, and H. Zhang, *Anxious or angry? effects of discrete emotions on the perceived helpfulness of online reviews.* MIS Quarterly, 2014. 38(2): p. 539-560.
[70] Zeng, X. and L. Wei, *Social ties and user content generation: Evidence from Flickr.* Information Systems Research, 2013. 24(1): p. 71-87.
[71] Zhang, Y., L. Feick, and V. Mittal, *How males and females differ in their likelihood of transmitting negative word of mouth.* Journal of Consumer Research, 2013. 40(6): p. 1097-1108.
[72] Zhu, F. and X. Zhang, *Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics.* Journal of Marketing, 2010. 74(2): p. 133-148.