Examining the Causality Loop between Online Reviews and Consumer Acquisition – a Granger Causality Study from YouTube

Full Paper

Jie Ren
Fordham University
jren11@fordham.edu

Introduction

People rely on online reviews to guide their purchasing behavior. According to industry research reports, a big proportion of consumers claim that their purchasing decision is largely influenced by online reviews (Deloitte 2007). In addition, consumers share these reviews with their family, friends or colleagues to make sense of their shopping experiences and deepen social connections (Peters and Kashima 2007). As a result, research increases significantly in size to address the impact of online reviews on consumer acquisition. Rather, this literature generates mixed results. For example, some studies showed that online review valence (from negative to positive) increases sales (e.g., Chevalier and Mayzlin 2006; Dellarocas, et al. 2007), while others showed that the driver of sales is the volume of online reviews, not the valence (e.g., Chen et al 2004; Liu 2006).

The confusion may come from a challenge embedded in online review systems (Duan et al. 2008): In these systems, people rely on online reviews; people also generate these same reviews. The network of products, reviews, and contributors is not only a reflection but also a driver of online behavior. However most of prior studies treated online reviews as exogenous (Chen et al. 2004; Chen et al. 2004; Dellarocas et al. 2007; Forman et al. 2008) and accordingly tested if online reviews drive sales. But it is equally important to verify if sales drive online reviews. In fact without considering the causality loop between online reviews and sales, some of the prior studies that are based on cross-sectional data may imply confusing results (Chatterjee 2001; Chen et al. 2004; Chen et al. 2004; Duan et al. 2008; Liu 2006). That is, high correlation between online reviews and sales in the cross-sectional settings can not only suggest online reviews drive sales, but also result from the causality in the other direction: sales drive online reviews.

Therefore, this paper aims to examine the dual influencer and indicator role of online reviews in relation to sales. By tracking consumer behaviors towards videos on YouTube for 64 days, this paper can provide evidences to reveal the “chicken-egg” problem in this literature.

The structure of this paper is as follows. The first part describes the literature about online reviews in relation to sales and discusses online review systems. The second part proposes a conceptual framework and hypotheses. The third part presents Granger Causality tests. The final part of this paper discusses the findings and especially the contributions to the online review literature.

Keywords

Comments, YouTube, Clicks, Causality, Review valence, Social media

Online Reviews as Exogenous to Sales
Recently, many researchers have studied the impact of online reviews, hoping to understand consumer behaviors and improve the designs of online review systems (e.g., Chevalier and Mayzlin 2006; Dellarocas et al. 2007; Ye et al. 2011). Generally, this literature considers online reviews as stimuli that trigger consumers’ decision-making. Positive reviews as stimuli are associated with pleasure; negative reviews with displeasure; review volume with awareness (e.g., Dellarocas, et al. 2007; Chevalier and Mayzlin 2006; Ren and Nickerson 2014; Sen and Lerman 2007; Ye et al 2009). Overall, the majority of this literature treated online reviews as an exogenous factor and some studies used cross-sectional data (Duan et al 2008).

To address the impact of online reviews, many scholars have examined online review valence (e.g., Dellarocas, et al. 2007; Chevalier and Mayzlin 2006; Sen and Lerman 2007; Ye et al 2009). For example, Chevalier and Mayzlin (2006) found a positive correlation between online review valence and online book sales. A different study found that online review volume and valence are both significant in forecasting movie sales (Dellarocas et al. 2007). Another study used Chinese online hotel booking data and found that online review valence affects the number of hotel rooms booked (Ye et al 2009). In addition, other papers related positive and negative reviews with product characteristics. Park and Lee (2009) found customers tend to avoid negatively-commented products more than they are attracted by positively-commented products. Furthermore, the impact is moderated by product category differences, with greater negativity-aversion observed in utilitarian products (e.g., durables) than in hedonic products (e.g., music, movies) (Sen and Lerman 2007).

By contrast, a minority of this literature suggested that people may not always exhibit aversion to negative online reviews and these reviews may not always hurt sales. For example, Berger et al (2010) showed that negative New York Times reviews increase sales for unknown books. Thelwall et al (2011) found that popular events on Twitter are normally associated with increases in negative valence strength. Also, Ghose and Ipeirotis (2011) includes text-mining analysis and found that when the review text is informative and detailed, products with negative online reviews are associated with increased sales.

To summarize, this valence-based approach has been employed on two levels. On the aggregate level, researchers analyzed the average ratings of online reviews and contrasted the effects of positive versus negative online reviews on product sales (Chevalier and Mayzlin 2006; Dellarocas et al 2007; Liu 2006; Ren and Nickerson 2013; Ye et al 2009). On the individual level, researchers, especially in the field of computer science, designed automated sentiment analysis system that extracts and summarizes the text of online reviews as positive or negative and focused on how the sentiment scores extracted from texts affect sales (e.g., Hu and Liu 2004; Pang et al. 2002).

In addition, to study the impact of online reviews, other papers focused on online review volume. They concluded that the driver of sales is the volume of online reviews, not the valence. In other words, essentially the awareness of products drives sales. For example Liu (2006) examined the box office performance of movies and claimed that online review volume can help increase consumers' awareness of movies and explain most of the variance in sales.

Therefore, the literature of online reviews is mixed: Focusing on different drivers (positive or negative reviews, review volume etc), the findings vary. Summarizing this literature, Duan et al. (2008) suggested that the confusion might come from two aspects. First, many researchers conducted their analysis using cross-sectional data. High correlation between online reviews and sales cannot indicate a one-way causality from online reviews to sales. Second, many studies treated online reviews as exogenous (Chen et al. 2004; Chen et al. 2004; Dellarocas et al. 2007; Forman et al. 2008). In online review systems, online reviews, however, don’t just drive sales; they are also the outcome of sales.

**Online Review Systems**

Online review systems are combinations of technologies and processes that collect online reviews and display them (e.g., Dellarocas 2003): A consumer reads online reviews to make a purchase decision; once a product is purchased, delivered and used, the consumer may evaluate such product and post a review (Figure 1).
On the one hand, consumers look for the aggregate information of online reviews (e.g., star ratings, review volumes and star rating distributions). They also evaluate individual reviews at a disaggregate level (e.g., Chen et al 2001; Mudambi and Schuff 2010). For example, more than one third of consumers will not book a hotel without reading review texts first (LateRooms 2012). Also, 90% of travelers avoid booking hotels described as “dirty” in online reviews (LateRooms 2012). At the disaggregate level individual review texts help customers to investigate the product more in depth before purchase.

On the other hand, after consumers place the orders and finish purchases, they upload their opinions onto online review systems, by either providing a star rating together with a text (Amazon), or voting for or against products (YouTube). For example, Dellarocas and Narayan (2006) have identified a few factors that affect consumers’ propensity to write a review about a movie.

In addition, there are debates about to what extent online reviews can represent consumers’ true consumption experiences (e.g., Koh et al. 2010; Hu et al. 2006; Hu et al. 2012). An example is given in Muchnik et al (2013)’s Science article. They found that reviewing behaviors are influenced by the valence of existing online reviews. Prior positive online reviews would increase the likelihood of positive ratings by 32%, whereas prior negative online reviews would trigger consumers to correct these negative online reviews.

In summary, online reviews play a dual role of influencer and indicator in relation to sales. Much has been discussed about the influencer role, but little regarding the indicator role. In fact, this dual role results from the design of online review systems. Understanding that consumers actively seek and upload online reviews, online systems use an online review display algorithm to present and aggregate reviews. Specifically, this algorithm sorts online reviews by a variety of review quality factors such as helpfulness, newness, valence and source credibility (Amazon, YouTube and Yelp). These reviews together with how they are collected, sorted and presented will influence other consumers in their purchase and/or reviewing decision-making (see Figure 1).

**Figure 1 Online Review Systems that Involve Consumer Participation**
Conceptual Framework and Hypotheses

As aforementioned, there is a causality loop between online reviews and consumer acquisition. Online reviewers serve as stimuli triggering consumers’ emotions towards the reviewed products. Purchase and consumption experiences in turn serve as antecedents of satisfaction or dissatisfaction towards the purchased products and can be described in both average star ratings and individual online review texts (Figure 2). Most of the recent work has focused on the former causality direction where online reviews drive sales. However, little attention has been spent on the latter direction where sales lead to online reviews. This paper examines both directions (Figure 2). By focusing on online review valence (positive or negative), this paper follows the majority of the online review literature (e.g., Dellarocas, et al. 2007; Chevalier and Mayzlin 2006; Sen and Lerman 2007; Ye et al 2009).

The Influencer Role and Indicator Role of Online Reviews

The influencer role of online reviews has been visited extensively in prior literature. Rather the majority has been motivated by the assumption: Positive online reviews increase sales, whereas negative online reviews hurt sales (e.g., Dellarocas, et al. 2007; Chevalier and Mayzlin 2006; Sen and Lerman 2007; Ye et al 2009). Intuitively, positive reviews provide information about satisfactory experiences with products and thus represent good product qualities. These reviews can help users filter out bad products and therefore help them find the right product more efficiently. Additionally, building on affective theories, positive emotional words expressed in online reviews can trigger the pleasure emotions of consumers who read these positive reviews (Ren and Nickerson 2014). So naturally, many prior works have found that positive online reviews drive sales. For example, Reichheld and Sasser (1990) indicated that positive information could increase revenue by attracting new customers.

Prior literature also studied how consumers write online reviews. However, the body of these studies is not as substantial as those that study the influencer role of online reviews. Moreover, these studies focused mainly on what motivates consumers to write reviews. For example, consumers write online reviews because of their desire for social interaction, desire for economic incentives, their concern for other consumers, and the potential to enhance their own self-worth (Hennig-Thurau, et al. 2004). These motivations to write online reviews also depend on their consumption experiences (e.g., Sundaram et al. 1998). For example, consumers write positive online reviews, as they want to relieve excitement caused by the purchase and use of products, or gain attention from helping others (e.g., Cheung and Lee 2012).
In addition, Huang et al. (2012) have linked consumers’ review reading behaviors with review writing behaviors. They found that after reading positive reviews, especially from friends or family, new consumers would assume the minimum quality of the associated product and are more likely to give positive reviews. So this paper aims to confirm these findings and propose the first hypothesis.

\[ H1 \text{ Positive online reviews cause consumer acquisition; consumer acquisition also causes positive online reviews.} \]

The causality loop also exists in regard to negative reviews. A few studies in the literature have suggested that negative reviews, on the other hand, may not always discourage consumer choice (e.g., Chmiel, et al., 2011; Homer and Yoon, 1992; Thelwall, et al, 2011). Evidences from negativity bias indicate the benefits of negative information: Negative information is more attention-grabbing in general and can cause more discussion than positive information (e.g., Chmiel, et al, 2011; Homer and Yoon, 1992; Thelwall, et al, 2011). Relating negativity bias, a minority of online review literature has found the positive impact of negative online reviews. But such positive impact exists with several contextual constraints. For example, negative online reviews increase sales when the review text is informative and detailed (Ghose and Ipeirotis 2011) or when products (e.g., books) are unknown (Berger et al 2010). In order to control for the contextual constrains and further test the impact of negative online reviews on sales in a general context, I aim to revisit these findings.

Additionally, the displeasure emotions triggered by negative online reviews can arouse consumers. For some consumers, protected by online anonymity and aroused by curiosity (Pinkerton and Zhou 2008; Zuckerman 1984), these displeasure emotions can lead to consumption behaviors.

The consumption behaviors related to negative valence can also contribute to review writing. For altruistic, anxiety reduction, vengeance, and advice seeking reasons, consumers write negative online reviews. That is, these consumers want to retaliate against the company associated with the negative consumption experiences, ease their anxiety caused by these experiences or help fellow consumers by warning them about these experiences (e.g., Sundaram et al. 1998). In addition, dissatisfied consumers are more likely to post online reviews than satisfied online reviews (Anderson 1998). This can be one reason that some celebrities seem to court controversy, leading to the cycle from negative word-of-mouth, to high awareness and purchases of their books, music albums or movies and further to negative word-of-mouth again and so on and so forth.

Therefore the second hypothesis relates to the causality loop regarding negative valence.

\[ H2 \text{ Negative online reviews cause consumer acquisition; consumer acquisition also causes negative online reviews.} \]

**Method**

I used YouTube data to test if online reviews drive sales and if there is a feedback causality loop from sales to online reviews. This platform has been used widely as a representative to study online review community (e.g., Burgess and Green 2013; Chintagunta et al. 2010): The updates in consumer feedback in YouTube are faster than those in other sites and within a short time period, behavioral patterns can be identified.

Following the majority of the literature on online reviews, I focused on one aspect of online reviews: valence. I tracked the top 25 recently-featured videos worldwide as identified by YouTube via YouTube API on a daily basis for 64 days. There are two reasons for this sampling. First, YouTube API can only return a list of no more than top 25 recently-featured videos. Second, recently-featured videos are more accessible to users and therefore would interact extensively with users. In order to control for the possible impact from product type, I also collected data from two Youtube categories: music (representing hedonic products) and education (representing utilitarian products). Also, in order to control for the possible impact of other events outside YouTube on YouTube videos, I increased the sample size. Specifically I
included data on 25 recently-featured music videos and 25 recently-featured education videos (Table 1). These videos were uploaded onto YouTube in January 2013.

<table>
<thead>
<tr>
<th>Examples of videos related to data collection</th>
<th>General</th>
<th>Music</th>
<th>Education</th>
</tr>
</thead>
<tbody>
<tr>
<td>PEOPLE ARE AWESOME 2013</td>
<td>Justin Timberlake - Suit &amp; Tie ft. JAY Z (MattyBRaps Cover)</td>
<td>Daniel Burrus: Predicting the Future</td>
<td></td>
</tr>
<tr>
<td>Child of the 90s</td>
<td>Internet Explorer</td>
<td>Let’s Get It On (Marvin Gaye Cover)</td>
<td>Aznar on Europe, America and Israel</td>
</tr>
<tr>
<td>Dramatic moment Prince Harry runs for his helicopter during Afghanistan interview</td>
<td>BoA 보아_그린 너 (Disturbance) Music Video</td>
<td>Fear and Oath-ing in D.C.</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Table 1 Titles of Video Examples</strong></th>
</tr>
</thead>
</table>

I collected data on these 75 videos on a daily basis from April 8th, 2013 to June 11th, 2013. The dataset includes the respective number of likes, dislikes and views of these videos (here I used the number of views to measure sales). Prior literature (Pavlou and Dimoka 2006) shows that users only read a small portion of individual online reviews and the overall valence score is the most influential aspect of a reviewing site. Therefore, I focused on the overall valence score. Also, similar to Muchnik et al. (2013) treating “vote for” and “vote against” as valence measure, I used dislikes and likes to measure online review valence.

Granger Causality tests were conducted relating valence and the number of views on the time-series data. In order to assume the stationarity, I only used daily data as the unit of analysis. Specifically, I used the daily number of views as the measure for customer acquisition per day. Even though YouTube users can click on each video multiple times, this is in fact similar to consumers’ purchase behaviors in other platforms including Amazon: consumer can buy the same products, especially consumer products, multiple times. In addition, I used the daily number of likes as the measure of positive online review valence per day, and the daily number of dislikes as the measure of negative online review valence per day.

**Results**

Using Granger Causality tests for the following bivariate VAR models with a time lag of 1 day (Table 2), I found that both positive online review valence and negative online review valence Granger-cause the popularity of YouTube videos (p < 0.01 for the former and the latter). Therefore H1 and H2 are supported.

However views only Granger-cause negative online review valence, but not positive valence (p < 0.05 for the former; p > 0.1 for the latter). Therefore, H4 is supported, but H3 is not.
Below summarizes whether each hypothesis is supported (Table 3).

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1 is partially supported.</td>
<td>H1 Positive online reviews cause consumer acquisition; but Consumer acquisition doesn’t cause positive online reviews.</td>
</tr>
<tr>
<td>H2 is supported.</td>
<td>H2 Negative online reviews cause consumer acquisition; consumer acquisition causes negative online reviews.</td>
</tr>
</tbody>
</table>

These results support the literature that positive online reviews lead to sales. These results also show the positive impact of negative online reviews on consumer acquisition. This particular finding relates to a few research that demonstrates that negative valence don’t necessarily hurt sales. In addition, these results indicate a feedback causality loop from sales to negative online reviews, but not to positive online reviews. This finding is surprising. First, it challenges the long-lasting assumption that sales is exogenous to online reviews and suggests caution to scholars that take such assumption (especially with positive online reviews). Second, it shows the moderating role of online review valence in such feedback causality loop. It seems that subjects were more likely to express their dislikes than their likes after watching a video. This particular finding can relate to the literature of negativity bias. These findings are discussed in details in the next section.

**Discussion**

This paper revisits the relationship between online reviews and consumer acquisition. It empirically identifies their causality loop and posts a challenge in prior online review literature. This challenge may have caused the mixed results about how online reviews affect on consumer acquisition. This is the first contribution to the literature.

This causality loop results from the design of online review systems: These systems collect, sort and display online reviews. That is, consumers reply on prior online reviews for purchases; after the purchase and use of products they write new online reviews for future consumers to read. Considering the ongoing process of reading and writing online reviews, the previous findings in the online reviews literature may not always hold. For example, capturing the interdependent relationship between online reviews and movie sales, Duan et al. (2008) refuted online review valence as the significant driver of movie sales in contrast to many of previous studies.

Because of the design of online review systems, it is intuitive to assume that online reviews drive sales and that sales in turn drive online reviews (Hypotheses 1 - 4). However, evidences were found that the causality loop doesn’t hold for all online reviews. In other words, negative online reviews drive and reflect on sales, whereas positive online reviews only drive sales. Identifying the moderating role of online review valence in the causality loop is the secondary contribution to the literature.

In addition to these two contributions, there are two interesting findings worth discussion. First, negative online reviews Granger-cause sales. Second, sales don’t Granger-cause positive reviews.
The Positive Impact of Negative Reviews on Customer Acquisition

This paper supports the previous findings on the positive impact of negative online reviews on sales. I can think of one plausible reason: these negative reviews can arouse consumers, provoke individuals’ curiosity and trigger individuals’ desire for entertainment. People are curious by nature (Berlyne 1954). Negative information is attention-grabbing: Especially, when negative online reviews describe arousal including anger, such curiosity and desire may be amplified. In fact, previous media research has indicated that the prevalent negative tone of television news stories is driven by people’s morbid curiosity: a mixture of compulsion, excitement, and fear. This is further supported by the popularity of fictional shows focused on macabre topics (Pinkerton and Zhou 2008; Zuckerman 1984). Maybe this is part of the reasons that some popular celebrities tend to court controversy: Controversy may spark negative online reviews.

This particular finding suggests caution for companies that aim to manage negative online reviews that they receive. Negative online reviews may not lead to sales reduction; they may increase sales in certain settings such as YouTube. Therefore in some settings or for certain products (e.g., hedonic products), a little controversy may be good.

Customer Acquisition not as an Antecedent of Positive Online Reviews

It is surprising to find that consumers don’t write positive online reviews, but write negative reviews after their consumption experiences in the YouTube setting. This finding is inconsistent with the pervasiveness of positive online reviews in online review systems and also inconsistent with the binomial and non-normal distribution of online review valence in Hu et al.’s (2006).

There may be two plausible reasons. First, consumers don’t necessarily express their satisfaction in online review texts right after the usage of products. That is, YouTube users may vote for a video whenever they want or even without watching the video. Second, YouTube users may be stingy in expressing their positive feelings for a video. This could explain the ubiquitous negative emotion carrying words in individual YouTube comments. This explanation is also aligned with the previous work that satisfied consumers are less likely to write online reviews than dissatisfied ones (Anderson 1998). Therefore the substantial amount of likes may not reflect the true reception in online review community. Rather these positive reviews or the likes may result from “water armies” who are hired by companies to manipulate public opinions. Future studies are needed to test these two reasons. But if the latter reason were verified, this paper would post a warning to online review systems/platforms such as Amazon. Actions are called for to delete fake positive reviews from these sites and re-establish the trust among users in online review communities.

Limitation and Outlook

This paper examines the causality loop in YouTube. However, there are three limitations about the choice of setting. First, YouTube only includes free products. Negative information triggers consumers’ curiosity and desire for entertainment (e.g., Ren and Nickerson 2014). Such curiosity and desire may be amplified for free products (Shampanier et al. 2007). Second, the products demonstrated on YouTube are only videos; most of the videos are slightly related to entertainment regardless of the category. Maybe because of the entertainment nature of YouTube videos, even with utilitarian products such as Education videos, negative online reviews would increase consumer acquisition in the YouTube setting. In typical online shopping communities such as YouTube, consumers face a much wider variety of products such as computers, desks, washers etc. And most of the products in these websites are priced. Therefore, future studies are needed to study other online review communities such as Amazon to generalize the findings of this paper on the causality loop. Third, this paper uses the extreme valences (likes and dislikes) to measure online review valence. Accordingly, the positive impacts of both positive and negative valences on consumer acquisition were identified. This in fact suggests for extreme valence reviews (5 star rating versus 1 star rating; likes versus dislikes), the driver of consumer acquisition is review volume. This finding is aligned with one stream of literature that studies movies industries. It also suggests that reviews for movies may have a bimodal distribution of review valence. Future studies should include reviews with 4 or 3 or 2 star ratings or treat valence as continuous rather than binary.
Conclusion

Online reviews drive sales, but also reflect on sales. This paper examined the causality loop between online reviews and sales and refuted the long-lasting assumption in the online review literature that online reviews are exogenous to sales. By doing so, this paper empirically tested Duan et al. (2008)’s reasons on the inconsistency in prior findings in the online review literature.

My findings show that the causality loop exists; however online review valence may moderate such causality loop. Specifically, negative online reviews play an influencer and indicator role in relation to sales; but positive online reviews only plays the influencer role. Therefore, these findings show caution to scholars that treat online reviews as exogenous to sales. Moreover, these findings draw attention to practitioners the positive impact of negative online reviews on consumer acquisition and the questionable sources of positive online reviews.

To summarize, by identifying the causality loop between online reviews and consumer acquisition, this paper hopes to contribute to the literature that studies consumer behavior in online review systems.
References


