Money Makes the Reviewer Go Round – Ambivalent Effects of Online Review Elicitation in B2B Markets

Completed Research

Jürgen Neumann[1]
Paderborn University
Juergen.Neumann@uni-paderborn.de

Dominik Gutt[1]
Paderborn University
Dominik.Gutt@uni-paderborn.de

Abstract

Online reviews have been thoroughly studied in business-to-consumer (B2C) markets, but less so in a business-to-business (B2B) context. On B2B review platforms like g2Crowd, both the platform and sellers heavily engage in review elicitation by inviting and incentivizing customers to write reviews. Despite the rich amount of research on online reviews, it is unclear how elicitation affects reviewing in B2B markets. Combining unique datasets from two B2B review platforms, we investigate different types of elicitation initiated by the platform and the seller, including email invitations and gift cards. Deriving hypotheses from relationship marketing and motivation theory, we use a difference-in-difference design to test our assumptions. Our results suggest that seller-initiated (platform-initiated) elicitation is associated with an increase (decrease) in review ratings, while review length is negatively associated with elicitation elicited by sellers, but not by platforms. Our results carry valuable managerial implications for B2B review platforms, sellers, and customers.

Keywords

Online Reviews, Review Elicitation, B2B Relationships, Third-Party Platform.

Introduction[2]

Online reviews are a cornerstone of today's e-commerce environment. Their main aim is to provide customers with reviews in order to reduce the information asymmetry between sellers and customers, and to establish trust in a product or service (Dellarocas 2003, Babic Rosario et al. 2016). However, research has found that reviews are often biased (Hu et al. 2017) and that different market participants may set different goals with respect to online review collection (Chen et al. 2017). Whereas review platforms such as Yelp and TripAdvisor might want to ensure that reviews of a particular hotel are truthful, the managers of that hotel seek to receive especially positive reviews to improve their own reputation. To allow both parties to achieve these goals, most online review systems implement features to elicit reviews from their customers, such as asking past customers for feedback via mail or offering payment in exchange for reviews.

For business-to-consumer (B2C) and consumer-to-consumer (C2C) markets, recent studies have analyzed the effect of elicitation measures on reviewing behavior such as asking for feedback via mail (Askalidis et al. 2017), offering rebates (Cabral and Li 2015) or introducing reviewer reputation systems (Shen et al., 2015). However, we are not aware of any study that has scrutinized the influence of review elicitation measures in a business-to-business (B2B) market, despite the suggestion in the budding stream of literature that B2B reviews appear to have an effect on sales, emphasizing the need to investigate online reviews in a

---

[1] This work has been partially supported by the German Research Foundation (DFG) within the Collaborative Research Center 901 "On-The-Fly Computing" under the project number 160364472-SFB901.

[2] A previous version of this study was published as Research-in-Progress, which differs substantially in terms of data, analysis and scope (Neumann and Gutt 2019).

B2B context (e.g., Gutt et al. 2019). Given the steady growth of B2B review platforms, such as TrustRadius, G2Crowd, and Capterra, research in this direction seems all the more timely. Such platforms are designed to enable customers to review business software in a professional context, and elicitation is used to a great extent as the majority of reviews are obtained by review elicitation. For instance, only about 28% of recent reviews for Skype for Business on G2Crowd were organic reviews (i.e. those spontaneously provided by customers), and the remainder were elicited either by mail or using monetary incentives. An interesting characteristic of B2B platforms is that reviews are invited or incentivized either by the platform itself or by the seller. Because eliciting reviews is costly in terms of the effort invested in obtaining sufficient reviews – and, of course, the cost of incentives – we need to understand whether and how such measures influence reviewing in the B2B market. Therefore, we pose the following research questions:

How do review elicitation measures by the seller and by the platform affect reviews in terms of rating and length in B2B markets?

To this end, we derive hypotheses from the literature on relationship marketing (Palmatier et al. 2006, Palmatier et al. 2009) and from insights on the motivation to write reviews (Askalidis et al. 2017). We hypothesize that, compared to organic reviews, reviews incentivized by the seller are more positive whereas reviews incentivized by the platform are more negative. Our intuition behind this lies in the fundamental difference between the outcomes pursued by sellers and by platforms, respectively, when incentivizing customers: sellers want high ratings whereas platforms prefer informative and truthful ratings. Finally, we hypothesize that an incentivized review is likely to be shorter because the reviewer's intrinsic motivation is crowded out by external incentives (Cerasoli et al. 2014). Analyzing a matched dataset of software reviews obtained from two different and large B2B review platforms, we find support for our hypotheses. In our data, we distinguish between reviews provided with our without invitation, as well as with or without an incentive – such as an Amazon gift card— by either the platform or the seller. We also observe whether the platform has offered to donate money to charity in exchange for writing a review. Conducting a difference-in-difference style regression with fixed effects suggests that, depending on the type of elicitation, the ratings of elicited reviews (measured on a scale from zero to ten) differ from organic reviews by about –0.029 to 0.69 rating units. We thus find evidence for our hypotheses, in that, compared to organic reviews, elicitation by a platform leads to lower ratings, and elicitation by the seller to higher ratings. As to the length of elicited reviews, we find that reviews elicited by the seller are substantially and significantly shorter than organic reviews.

To the best of our knowledge, this is the first study to empirically scrutinize the impact of different types of elicitation that involves (1) a B2B review context and (2) elicitation measures from both a platform and a seller. Our results carry substantial managerial implications. Platforms need to be aware that, although eliciting reviews can lead to negative/more truthful reviews, the resulting reviews are significantly shorter, which is often associated with a lower perceived helpfulness (Kuan et al. 2015). Sellers, however, seem to be able to increase their reputation by giving monetary rewards to customers in exchange for their reviews. From a customer perspective, these insights are valuable because, when basing their purchase decisions on reviews, they should be aware of the differences between organic and elicited reviews, and the potential biases induced by different forms of elicitation. Finally, we enrich the scholarly discussion of online reviews in B2B markets – especially on its drivers – that has been largely neglected by prior research. Moreover, our results underscore the need for research in B2C markets to investigate whether reviews elicited by sellers or by platforms similarly follow this pattern.

Related Literature

Trust between two business parties is one of the main facilitators of business transactions (McKnight et al. 2002). The literature has, by and large, reached the consensus that online reviews are a primary way to build trust between customers and sellers in business-to-consumer (B2C) markets (Ba and Pavlou 2002, Pavlou and Dimoka 2006). In essence, online reviews help erode the information asymmetry between buyer and seller prior to purchase, allowing the buyer to inspect features of the products and characteristics of the sellers, based on the feedback provided by previous customers (Dellarocas 2003). Studies have thus documented a wide range of positive implications of online reviews. For instance, they can support customer decision making, and positive online reviews can increase a seller's sales (Babic Rosario et al. 2016). Due to these effects, a considerable number of studies have investigated how sellers in B2C markets can effectively elicit online reviews from buyers.

Scholars have analyzed multiple methods of review elicitation and their respective impact on a review's rating and length. Table 1 provides an overview of existing studies in this specific research area – including our own to show how it adds to this particular stream of literature. Please note that, while we are aware that a number of studies investigate the impact of review elicitation on the likelihood to write a review (Chen, 2010), for the purpose of this study, we focus solely on those that investigate the length and the rating of a review.

Study	**Types of Elicitation**	**Source of Elicitation**	**Data Source**	**Dependent Variable**	**Market Type**
Askalidis et al. (2017)	Mail invitation	Seller	4 major online retailers	Rating, length, helpfulness	B2C
Burtch et al. (2018)	SMS invitation with/without coupon or social norm	Seller	Retailer, Amazon Turk	Volume, length	B2C
Cabral and Li (2015)	Rebate	Seller	eBay	Rating	B2C/C2C
Chen et al. (2017)	Rebate	Seller, Platform	None (theoretical model)	Truthfulness/rating, volume	C2C
Khern-am-nuai et al. (2018)	Coupon	Seller	Retailer, Amazon	Rating, length, textual features	B2C
Shen et al. (2015)	Reviewer reputation system	Seller	Barnes andNoble, Amazon	Average rating deviation, volume	B2C
This study	Mail invitation with/without gift card or donation to charity	Seller, Platform	Two B2B review platforms	Rating, length	B2B

Table 1. Overview of Research on Review Elicitation.

By and large, the elicitation types investigated in the literature can be categorized as follows: (1) no incentive, (2) non-monetary incentive, and (3) monetary incentive. The first category comprises studies that analyze email or SMS invitations without incentives. Empirical evidence suggests that reviewers who are invited by email write more positive reviews and finds that these are both shorter and perceived as less helpful (Askalidis et al. 2017). In the second category, instead of simply asking for reviews, studies have analyzed the effect of incorporating social norms into elicitation mails, such as comparing the customer's reviewing activity with that of other reviewers. It has been found that emphasizing social norms positively affects the length of reviews (Burtch et al. 2018). Scholars have also studied the effects of non-monetary incentives, such as reviewer reputation programs which give reviewers rankings (Shen et al. 2015). In the third category, studies investigated methods that include giving reviewers money in the form of coupons or rebates. There is a positive relationship between offering rebates or coupons in exchange for reviews and the rating of reviews (Cabral and Li 2015, Khern-am-nuai et al. 2018), and one theoretical model suggests that if a platform awards the rebate, reviews tend to be more negative/truthful (Chen et al. 2017).

Our study adds to this stream of literature but also differs from previous literature in two major ways. First, to the best of our knowledge, we are the first to analyze review elicitation in a B2B context. Second, most of the literature focuses on elicitation initiated by the seller. Our study provides a unique setting in which we can observe elicitations initiated by two different sources (sellers and platform) across two different platforms.

Theoretical Background and Hypotheses

To delineate hypotheses on the effective elicitation of online reviews in B2B markets, we borrow theory from relationship marketing, defined as “all marketing activities directed towards establishing, developing, and maintaining successful relational exchanges” (Morgan and Hunt 1994, p. 22). A sizeable body of literature on relationship marketing has found that when one party invests in the business-customer relationship, it can spur favorable reactions from the other because it prompts gratitude-related reciprocal behavior by the recipient of the investment (Palmatier et al. 2009). Sellers, for example, benefit from such investments in the form of customer loyalty (Garnefeld et al. 2013), commitment (Palmatier et al. 2006), and word of mouth (Palmatier et al. 2006). This study is mainly interested in the relationship between relationship investment and electronic word of mouth—that is, online reviews, which is the most prevalent form of word of mouth in online markets (Dellarocas 2003, Hennig-Thurau et al. 2004). According to relationship marketing literature, investments generate expectations of reciprocity that lead to favorable word of mouth dissemination by the customer about the party that invested in the relationship (Anderson and Weitz 1989, Ganesan 1994). A recent study on Airbnb, for example, provides empirical evidence for the positive impact of seller effort on the seller’s rating through reciprocal buyer behavior (Proserpio et al. 2018). The literature identifies relationship investments into the relationship with a customer in the form of gifts, sales agent support, or loyalty programs, for instance (Wulf et al. 2001, Ganesan 1994). In our case, relationship investment is represented by the seller eliciting a review. Depending on the type of elicitation, the investment differs in strength. Providing a gift in the form of an Amazon gift card should be a stronger investment than kindly asking for a review without further incentives. Because relationship investment by a seller should prompt positive gratitude-based reciprocity, we expect that eliciting a review would lead to higher ratings for the seller. Thus, we formulate Hypothesis 1 as follows:

Hypothesis 1: A customer’s online rating should be higher than an organic rating when the seller elicits the review.

However, in our environment, it is not just sellers who actively elicit online reviews from customers; the platform itself can elicit reviews on the software that a customer has bought from a seller. Again, relationship marketing states that customers react favorably to the party that invested in the relation-ship. Because the party in question is the platform, not the seller, and the platform is interested in truthful and informative reviews on a product or service, it follows that a customer would not necessarily provide a high online rating. In line with this, Chen et al. (2017) theorize that when platforms ask for reviews, they should receive truthful and informative ones. When a platform elicits a review, it openly communicates that it would like to receive informative reviews. Thus, reviewers are encouraged to be critical (that is, rather negative [Sen and Lerman, 2007]). Further, because they are not directly asked by the seller, they should be less inclined to support the seller with a particularly high rating. Therefore, a review elicited by the platform should have a lower rating than a review elicited by the seller. Previous studies have demonstrated that organic reviews are prone to the positivity bias because reviewers are more likely to write a review if they had a positive experience (Hu et al. 2017). If customers are now asked to write reviews for products that they would not review otherwise, the positivity bias should be alleviated. To summarize, customers asked by the platform should be more critical, and their ratings should not be driven by a positivity bias, which implies ratings that are lower than organic ones. We formulate Hypothesis 2:

Hypothesis 2: A customer’s online rating should be lower than an organic rating when the platform elicits the review.

Finally, when a party engages in review elicitation, it tries to extrinsically motivate the reviewer. The phenomenon that extrinsic incentives substitute intrinsic motivation has been studied extensively by psychology (Cerasoli et al. 2014). Similarly to other findings on extrinsic motivation for public goods contribution, the literature has recognized that review elicitation provides extrinsic stimuli that crowd out intrinsic motivation of review writing (Askalidis et al. 2017, Khern-am-nuai et al. 2018). Incentives provided by the seller or the platform are not tied to review length, resulting in reviewers spending less effort in writing longer reviews because they are driven by extrinsic stimuli. Thus, we formulate Hypothesis 3 as follows:

Hypothesis 3: A customer’s online review should be shorter than an organic rating when the review is elicited regardless of who elicits the review.

Empirical Analysis

Data

Our comprehensive dataset consists of a complete review history from two distinct B2B reviewing platforms. On both platforms, consumers can write reviews on business software such as Skype for Business or Shopify, and give a rating on a scale from 1 to 10. The total number of reviews accumulated between August 2012 and August 2018, on both platforms taken together, is 500,926. To prepare the data according to our identification strategy, we identify software products that have been reviewed on both platforms, which resulted in 1,124 products. The total number of reviews for these is 236,104. Table 2 presents their descriptive statistics.

	Platform A				Platform B			
Variable	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
RATING	8.5804	1.7293	1	10	8.5421	1.6675	1	10
LENGTH	729.148	427.9284	0	11451	2149.063	1155.304	209	13610
ANONYMOUS	.5072	.4999	0	1	.3238	.4680	0	1
SELLER_ INVITE	.0004	.0195	0	1	.1812	.3852	0	1
SELLER_ INCENTIVE	.0079	.0887	0	1	.0876	.2828	0	1
PLATFORM_ INCENTIVE	.6658	.4717	0	1	.1703	.3759	0	1
PLATFORM_ INVITE	.0288	.1672	0	1	.0764	.2657	0	1
PLATFORM_ CHARITY	.0162	.1262	0	1	.	.	.	.
PLATFORM_ RAAS_ INVITE	.	.	.	.	.0009	.0293	0	1
PLATFORM_ RAAS_ INCENTIVE	.	.	.	.	.2680	.4429	0	1
Number of Observations	213,967				22,137			

Table 2. Descriptive Statistics

For each review, our dataset contains the review text from which we calculate its length in terms of number of characters ($LENGTH$), a rating ($RATING$) and information on the review author (i.e., industry, company size, etc.) as well as a unique identifier for the author unless they decided to stay anonymous ($ANONYMOUS$). Moreover, we observe the date of the review's publication. For both platforms, we also observe whether a review was elicited by the seller, by the platform, or whether it was given organically, i.e. without elicitation. Sellers can either invite customers to write a review without providing an incentive ($SELLER\_INVITE$) or offer an incentive in form of an Amazon Gift Card ($SELLER\_INCENTIVE$). Similarly, the platforms can also send out invitations with or without an incentive, respectively ($PLATFORM\_INVITE$, $PLATFORM\_INCENTIVE$). Additionally, one of the platforms sometimes offers to donate to a good cause in return for writing a review ($PLATFORM\_CHARITY$). The other platform offers a "Review as a Service"-Option which can be booked by a seller. In this case, the platform invites or incentivizes consumers on behalf of the seller ($PLATFORM\_RAAS\_INVITE$, $PLATFORM\_RAAS\_INCENTIVE$). We summarize the variables $SELLER\_INVITE$ and $SELLER\_INCENTIVE$ in a single variable called $SELLER\_ELICITED\_COUNT$ to capture seller-initiated elicitation. We use the remaining elicitation-related variables to form the variable $PLATFORM\_ELICITED\_COUNT$ to describe platform-initiated elicitation.

Empirical Model and Identification Strategy

We aggregate our data on the software product level per quarter so that, for each product, we are able to compare both their quarterly average ratings and their quarterly average review text lengths (Y_{ijt}) between the two platforms. The quarterly average rating and the quarterly average review text length is given on platform $i \in \{A, B\}$ for software j at quarter t as represented in the following function:

$$Y_{ijt} = \beta_0 + \beta_1 SELLER\_ELICITED_{ijt} + \beta_2 PLATFORM\_ELICITED_{ijt} + \alpha_{jt} + \delta_{ij} + \mu_{it} + \epsilon_{ijt} \quad (1)$$

Our main explanatory variables of interest are $SELLER\_ELICITED\_COUNT_{ijt}$ and $PLATFORM\_ELICITED\_COUNT_{ijt}$. They describe the number of reviews elicited by the seller or the platform, respectively. δ_{ij} describes platform-software fixed effects. For instance, one review platform could mainly attract reviewers of a certain type of software (e.g., for email marketing) whereas the other mostly attracts reviewers of a different type of software (e.g., big data analytics tools). δ_{ij} controls for these unobserved differences across review platforms. Moreover, μ_{it} denotes the general platform-specific time trend. For example, reviewers on the same platform could become more positive over time. α_{jt} denotes unobserved quarterly changes on the software-level. These include changes in the software's quality and the software seller's marketing or sales strategy that could impact the software's rating. Finally, the term ϵ_{ijt} captures any remaining factors that could influence the software's quarterly average rating.

Even though the platform-software fixed effects δ_{ij} and the quarterly fixed effects μ_{it} in Equation (1) control for a fair amount of unobserved heterogeneity, changes in the software quality, the seller's marketing or sales strategy (captured by α_{jt}) could still bias the relationship between the rating of a reviewer and the elicitation activity of sellers and platforms. Following the approach established by Chevalier and Mayzlin (2006) and Zhu and Zhang (2010), we pursue the strategy to difference out α_{jt} to mitigate this bias. The main idea behind this identification strategy is that any unobserved changes to α_{jt} should affect both platforms. Thus, calculating the difference across platforms between our dependent variables and our right-hand side variables should remove this bias. The resulting model is displayed in Equation (2).

$$\begin{aligned} Y_{Ajt} - Y_{Bjt} = \gamma_0 &+ \gamma_1 \left(SELLER\_ELICITED\_COUNT_{Ajt} - SELLER\_ELICITED\_COUNT_{Bjt}\right) \\ &+ \gamma_2 \left(PLATFORM\_ELICITED\_COUNT_{Ajt} - PLATFORM\_ELICITED\_COUNT_{Bjt}\right) \\ &+ \left(\delta_{Aj} - \delta_{Bj}\right) + \left(\mu_{At} - \mu_{Bt}\right) + \left(\epsilon_{Ajt} - \epsilon_{Bjt}\right) \end{aligned} \quad (2)$$

Removing software level changes over time from the equation yields two additional benefits. First, by introducing software-level fixed effects $\theta_j = \left(\delta_{Aj} - \delta_{Bj}\right)$, we can control for time constant differences between the platforms for each individual software. For example, a software seller could allocate different marketing budgets for review elicitation for the two platforms. As long as the level of the differences are constant for a quarter, θ_j controls for these differences. Second, by denoting $\nu_t = (\mu_{At} - \mu_{Bt})$ as time-fixed effects, we can control for the differences in time trends between the two platforms. For instance, one platform could be more successful in attracting new reviewers in general compared to the other platform, leading to a disproportionate growth in the reviewer population between platforms. ν_t controls for these differences. We let $\eta_{ijt} = \left(\epsilon_{Ajt} - \epsilon_{Bjt}\right)$ be the unobserved error term that describes the remaining software-specific differences over time between the two platforms. Finally, to account for differences in software popularity, we include the quarterly review count of each platform as a control variable (denoted by

$COUNT$). Thus, we conduct an Ordinary Least Squares regression with fixed effects to estimate the following equation.

$$\Delta Y_{jt} = \gamma_0 + \gamma_1 \Delta SELLER\_ELICITED\_COUNT_{jt} + \gamma_2 \Delta PLATFORM\_ELICITED\_COUNT_{jt} + \gamma_3 \Delta COUNT_{jt} + \theta_j + \nu_t + \eta_{ijt} \quad (3)$$

Results

Table 3 presents the results of our baseline model (Equation 3) with ΔAVG_RATING in column (1) and ΔAVG_LENGTH in column (2) as the dependent variable.

Dependent Variable	(1) ΔAVG_RATING	(2) ΔAVG_LENGTH	(3) ΔAVG_RATING	(4) ΔAVG_LENGTH
$\Delta SELLER\_ELICITED\_COUNT$	.0082* (.0045)	-7.0229*** (2.6341)	.	.
$\Delta PLATFORM\_ELICITED\_COUNT$	-.0034572** (.0017)	.8711 (1.1322)	.	.
$\Delta COUNT$	.0035** (.0015)	-.7025 (1.0113)	.0035** (.0015)	-.7010 (1.0114)
$\Delta SELLER\_INVITE$	.	.	.0027 (.0073)	-14.4447*** (4.3062)
$\Delta SELLER\_INCENTIVE$	.	.	.0137* (.0080)	.0073 (4.083725)
$\Delta PLATFORM\_INVITE$	.	.	.0093 (.0128)	9.9898 (6.7097)
$\Delta PLATFORM\_INCENTIVE$	.	.	-.0035** (.0017)	.8013 (1.1370)
$\Delta PLATFORM\_CHARITY$	.	.	-.0029 (.0051)	.8142 (2.4064)
$\Delta PLATFORM\_RAAS\_INCENTIVE$	.	.	-.0070** (.0035)	-2.2389 (1.8572)
$\Delta PLATFORM\_RAAS\_INVITE$	.	.	-.0642 (.1047)	46.8473 (83.9707)
Software-Fixed Effects	✓	✓	✓	✓
Time-Fixed Effects	✓	✓	✓	✓
N	4,171	4,171	4,171	4,171
Note: Robust standard errors in parentheses. *<p.1; **<p.05; ***<p.01.				

Table 3. Baseline Results

Our results suggest that if the difference between seller-elicited reviews between the two platforms increases by 1 review, this is associated with a statistically significant increase in the quarterly average rating by 0.01 points. Put differently, with every 25 reviews elicited by a seller, they gain a ¼-star increase in their quarterly average rating. However, for platform-elicited reviews, we find a statistically significant decrease in the quarterly average rating by 0.003 points. Therefore, platforms would need to elicit approx. 83 reviews in order to decrease a product's quarterly average rating by 1/4 stars. In sum, we find support for Hypothesis 1 and 2. Regarding review length, we find that a seller-elicited rating is associated with a reduction of the quarterly average review length by seven characters. In contrast, for platform-elicited reviews, our results do not indicate a statistically significant relationship between the number of reviews and review length. This suggests that platforms can effectively elicit reviews that are more negative (i.e. more likely to be perceived as useful) than organic or seller-elicited reviews, without this resulting in shorter (i.e. less informative) reviews. This supports our intuition that platforms can elicit more informative reviews without loss in length, whereas seller-elicited reviews are significantly shorter on average. Thus, we only find partial support for Hypothesis 3, as seller-elicited reviews are shorter, whereas platform-elicited reviews are not.

Splitting up the independent variables into their individual components (Column (3) and (4)) yields further insights. In line with our theoretical assumption, our results are driven by monetary incentives offered by the seller, as they are associated with an increase in ratings. There seems to be no statistically significant relationship for neither plain review invitations (i.e. those without an incentive) nor donations to charity, in relation to any of our dependent variables (quarterly average rating and length). The only exception is that invitations by the seller without monetary compensation are associated with a decrease in length.

Further Analysis and Robustness Checks

While our baseline model effectively mitigates potential sources of endogeneity, it does not consider reviewer heterogeneity, assuming that reviewers are homogenous across platforms. If particularly positive or negative reviewers are more likely to be invited/incentivized to review specific software on one platform compared to the other, this may confound our result. To address this confounding factor, we take advantage of the fact that we can observe author IDs on a review level and estimate the following equation in an OLS regression with fixed effects.

$$Y_{ijrt} = \beta_0 + \beta_1 SELLER\_ELICITED_{ijrt} + \beta_2 PLATFORM\_ELICITED_{ijrt} + \beta_3 \gamma_{ijr} + \alpha_{ij} + \delta_{ir} + \mu_{it} + \epsilon_{ijrt} \tag{1}$$

The variable Y_{ijt} represents our outcome variable ($RATING$, $LENGTH$) of a review for software j written by reviewer r at time t on platform i. $SELLER\_ELICITED_{ijrt}$ and $PLATFORM\_ELICITED_{ijrt}$ are binary variables that indicate whether the review was elicited by the seller or the platform, respectively. γ_{ijrt} represents a vector of review-related control variables such as the average rating and count of all reviews prior to the current one. $\alpha_{ij}, \delta_{ir}, \mu_{it}$ describe software-, reviewer- and time-level fixed effects per platform, respectively. ϵ_{ijrt} is the random unobserved error-term. We include all reviews from our initial (500,926 reviews) dataset into this analysis as incorporating reviewer level fixed effects demands more variation within a reviewer's review history[3]. Note that the actual number of observations is 136,653, which is smaller than 500,926 as single observations with respect to the fixed effects drop out. The estimated coefficients for RATING as a dependent variable are as follows: .6908 for SELLER_ELICITED and -.0293 for PLATFORM_ELICITED. Using LENGTH as a dependent variable, the coefficients are as follows: -226.4016 for SELLER_ELICITED and -21.8664 for platform elicited. One coefficient is statistically significant at the 10%-level, whereas all others are significant at the 1%-level. These results support our hypotheses and our baseline model, reassuring us that our baseline results are not driven by unobserved reviewer heterogeneity. The results also demonstrate that the magnitude of the relationship between elicitation types and ratings is much more pronounced on the individual rating level than on the quarterly average rating level. Seller-elicited ratings are, for instance, more than half a rating point higher than organic ratings, which is quite substantial.

We conduct further checks to ensure the robustness of our results. First, we restrict our sample only to those observations with more than one review on both platforms to ensure that our results are not driven by a low reviewing activity on one of the platforms. The results remain qualitatively unchanged. Second, we chose quarters as the time frame of aggregation which might be too narrow or too loose. We re-aggregate our raw data on a yearly and monthly level and re-estimate our model, again with similar results.

[3] Restricting this analysis to the sample of our baseline model yields similar results.

Conclusion

The literature on online reviews has experienced a tremendous increase in studies focusing on the economic effects of reviews and the factors that drive their generation. Yet, these studies have been confined to the domain of B2C and C2C markets, almost exclusively neglecting B2B markets. Our paper is, to the best of our knowledge, among the first to narrow this gap. Our study focuses on two B2B online review platforms and investigates the ways online reviews are elicited on these platforms. While previous empirical studies on review elicitation in B2C markets have exclusively focused on review elicitation by sellers, we examine review elicitation that can be either initiated by the seller or by the platform. Because these two parties pursue different aims in eliciting reviews – sellers want positive reviews and platforms want informative and truthful reviews – we hypothesize differential effects of review elicitation on the rating magnitude given by reviewers. Our empirical evidence gleaned from a matched sample of reviews for professional software from two B2B reviewing platforms presents support for our hypotheses. Our further analyses and robustness checks provide additional evidence that our results are not driven by unobserved reviewer heterogeneity and that our effects are most pronounced for elicitation measures that involve monetary gifts.

Our results contribute to scholarly understanding of online reviews in two ways. First, this study is among the first to examine review elicitation in B2B markets, providing evidence for the effectiveness of gifts in eliciting reviews from business customers. Second, we extend the perspective of review elicitation from a seller-customer relationship by adding a third party, namely the review platform, whose desired outcome in eliciting reviews differs to that of sellers. In doing so, we highlight differential effects of review elicitation on the magnitude of ratings, depending on which party initiates the elicitation. In this way, we lend empirical support to the theoretical study by Chen et al. (2017), who have formulated differential effects of review elicitation using an analytical model. These results carry implications for both scholars and practitioners. First, researchers investigating review elicitation measures should clearly differentiate between the parties that initiate review elicitation. Second, our results could represent a candidate solution to the phenomenon of "rating inflation" (Filippas et al. 2018, Zervas et al. 2015), where online ratings become uninformative because all ratings in a market are very high. Rating elicitation by the platform could help counteract "rating inflation" as it increases the informational value of a review and decreases the magnitude of ratings. Our results are also relevant to practitioners. Platforms could implement review elicitation to acquire informative and truthful reviews, in order to keep them appealing to users looking for reviews on their website. In particular, platforms might elicit reviews from products whose sellers are also actively eliciting reviews, in order to prevent an overly positive average review for this product. In order to improve their average ratings, sellers should invest money in building a relationship with their customers, e.g. in the form of monetary incentives. Finally, our results open up new avenues for future research. Future studies could investigate whether differential effects between seller- and platform elicited reviews also exist in B2C or C2C markets. Moreover, future research could investigate the economic effects of reviews on sales in B2B markets, for instance, which is beyond the scope of our study.

REFERENCES

Anderson, E., and Weitz, B. 1989. "Determinants of Continuity in Conventional Industrial Channel Dyads," *Marketing Science* (8:4), pp. 310–323.

Askalidis, G., Kim, S. J., and Malthouse, E. C. 2017. "Understanding and Overcoming Biases in Online Review Systems," *Decision Support Systems* (97), pp. 23–30.

Ba, S., and Pavlou, P. A. 2002. "Evidence of the Effect of Trust Building Technology in Electronic Markets: Price Premiums and Buyer Behavior," *MIS Quarterly* (26:3), pp. 243–268.

Babic Rosario, A., Sotgiu, F., Valck, K. de, and Bijmolt, T. H.A. 2016. "The Effect of Electronic Word of Mouth on Sales: A Meta-Analytic Review of Platform, Product, and Metric Factors," *Journal of Marketing Research* (53:3), pp. 297–318.

Burtch, G., Hong, Y., Bapna, R., and Griskevicius, V. 2018. "Stimulating Online Reviews by Combining Financial Incentives and Social Norms," *Management Science* (64:5), pp. 1975–2471.

Cabral, L., and Li, L. 2015. "A Dollar for Your Thoughts: Feedback-Conditional Rebates on eBay," *Management Science* (61:9), pp. 2052–2063.

Cerasoli, C. P., Nicklin, J. M., and Ford, M. T. 2014. "Intrinsic Motivation and Extrinsic Incentives Jointly Predict Performance: A 40-Year Meta-Analysis," *Psychological Bulletin* (140:4), pp. 980–1008.

Chen, L., Jiang, T., Li, W., Geng, S., and Hussain, S. 2017. "Who Should Pay for Online Reviews?: Design of an Online User Feedback Mechanism," *Electronic Commerce Research and Applications* (23), pp. 38–44.

Chevalier, J. A., and Mayzlin, D. 2006. "The Effect of Word of Mouth on Sales: Online Book Reviews," *Journal of Marketing Research* (43:3), pp. 345–354.

Dellarocas, C. 2003. "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms," *Management Science* (49:10), pp. 1407–1424.

Filippas, A., Horton, J. J., and Golden, J. 2018. Reputation Inflation. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3136473.

Ganesan, S. 1994. "Determinants of Long-Term Orientation in Buyer-Seller Relationships," *Journal of Marketing* (58:2), p. 1.

Garnefeld, I., Eggert, A., Helm, S. V., and Tax, S. S. 2013. "Growing Existing Customers' Revenue Streams through Customer Referral Programs," *Journal of Marketing* (77:4), pp. 17–32.

Gutt, D., Neumann, J., Zimmermann, S., Kundisch, D., and Chen, J. 2019. "Design of review systems – A strategic instrument to shape online reviewing behavior and economic outcomes," *The Journal of Strategic Information Systems* (Forthcoming).

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., and Gremler, D. D. 2004. "Electronic Word-of-Mouth via Consumer-Opinion Platforms: What Motivates Consumers to Articulate Themselves on the Internet?" *Journal of Interactive Marketing* (18:1), pp. 38–52.

Hu, N., Pavlou, P. A., and Zhang, J. 2017. "On Self-Selection in Online Product Reviews," *MIS Quarterly* (41:2), pp. 449–471.

Khern-am-nuai, W., Kannan, K., and Ghasemkhani, H. 2018. "Extrinsic versus Intrinsic Rewards for Contributing Reviews in an Online Platform," *Information Systems Research* (29:4), pp. .779–1068.

Kuan, K. K. Y., Hui, K.-L., Prasarnphanich, P., and Lai, H.-Y. 2015. "What Makes a Review Voted? An Empirical Investigation of Review Voting in Online Review Systems," *Journal of the Association for Information Systems* (16:1), pp. 48–71.

McKnight, D.H., Choudhury, V., and Kacmar, C. 2002. "The Impact of Initial Consumer Trust on Intentions to Transact with a Web Site: A Trust Building Model," *The Journal of Strategic Information Systems* (11:3-4), pp. 297–323.

Morgan, R. M., and Hunt, S. D. 1994. "The Commitment-Trust Theory of Relationship Marketing," *Journal of Marketing* (58:3), p. 20.

Neumann, J., Gutt, D. 2019. "He Who Pays the Piper Calls the Tune: Online Review Elicitation by Sellers and Third-Party Platforms in B2B Markets", in: *Proceedings of the Twenty-Seventh European Conference on Information Systems (ECIS)*, Research-in-Progress, Stockholm, Schweden (forthcoming).

Palmatier, R. W., Dant, R. P., Grewal, D., and Evans, K. R. 2006. "Factors Influencing the Effectiveness of Relationship Marketing: A Meta-Analysis," *Journal of Marketing* (70:4), pp. 136–153.

Palmatier, R. W., Jarvis, C. B., Bechkoff, J. R., and Kardes, F. R. 2009. "The Role of Customer Gratitude in Relationship Marketing," *Journal of Marketing* (73:5), pp. 1–18.

Pavlou, P. A., and Dimoka, A. 2006. "The Nature and Role of Feedback Text Comments in Online Marketplaces: Implications for Trust Building, Price Premiums, and Seller Differentiation," *Information Systems Research* (17:4), pp. 392–414.

Proserpio, D., Xu, W., and Zervas, G. 2018. "You Get What You Give: Theory and Evidence of Reciprocity in the Sharing Economy," *Quantitative Marketing Economics* (16:4), pp. 371–407.

Sen, S., and Lerman, D. 2007. "Why are you telling me this?: An Examination into Negative Consumer Reviews on the Web," *Journal of Interactive Marketing* (21:4), pp. 76–94.

Shen, W., Hu, Y., and Ulmer, J. R. 2015. "Competing for Attention: An Empirical Study of Online Reviewers' Strategic Behavior," *MIS Quarterly* (39:3), pp. 683–696.

Wulf, K. de, Odekerken-Schröder, G., and Iacobucci, D. 2001. "Investments in Consumer Relationships: A Cross-Country and Cross-Industry Exploration," *Journal of Marketing* (65:4), pp. 33–50.

Zervas, G., Proserpio, D., and Byers, J. 2015. A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2554500.

Zhu, F., and Zhang, X. 2010. "Impact of Online Consumer Reviews on Sales: The Moderating Role of Product and Consumer Characteristics," *Journal of Marketing* (74:2), pp. 133–148.