IN THE EYE OF THE BEHOLDER? EMPIRICALLY DECOMPOSING DIFFERENT ECONOMIC IMPLICATIONS OF THE ONLINE RATING VARIANCE

Research paper

Gutt, Dominik[1], Paderborn University, Paderborn, Germany, dominik.gutt@upb.de

Abstract

The growing body of literature on online ratings has reached a consensus of the positive impact of the average rating and the number of ratings on economic outcomes. Yet, little is known about the economic implication of the variance of online ratings, and existing studies have presented contradictory results. Therefore, this study examines the relationship between the variance of online ratings and the price and sales for digital cameras from Amazon.com. The key feature of our study is that we employ and validate a machine learning approach to decompose the online rating variance into a product failure-related and taste-related share. In line with our theoretical foundation, our empirical results highlight that the failure-related variance share is negatively associated with price and sales, and the taste-related share exhibits a positive relationship with price and sales. Our results highlight a new perspective on the online rating variance that has been largely neglected by prior studies. Sellers can benefit from our results by adjusting their pricing strategy and improving their sales forecasts. Review platforms can facilitate the identification of product failure-related ratings to support the purchasing decision process of customers.

Keywords: Online Rating Variance, Text Mining, Econometrics, User-Generated Social Media.

1 Introduction

In the era of digitization, user-generated social media content, such as online ratings, are a primary way to acquire online information about goods and services. Online ratings are a driving force behind online (Cabral and Hortaçsu 2010) as well as offline (Anderson and Magruder 2012) customer behavior because they facilitate the evaluation and comparison of the quality of a product or service. In fact, surveys report that 90% of all online purchase decisions are influenced by online ratings (Drewnicki 2013), and they are considered a crucial factor of success in Amazon's business (Allen 2016).[2] Unsurprisingly, researchers have been keen on investigating the impact of this information on customer behavior. As a result, the positive effects of the average rating (Luca 2016; Anderson and Magruder 2012; Li and Hitt 2008) and of the number of ratings (Chevalier and Mayzlin 2006; Duan et al. 2008) are well established in the literature. Also, numerous studies have investigated the drivers behind the often-observed J-shape of the rating distribution (Hu et al. 2017; Koh et al. 2010).

Although previous studies generally depict consistent results concerning the impact of the average rating and the number of ratings, the effect of the variance of online ratings on economic outcomes remains insufficiently understood. The variance of online ratings is a metric indicating to what extent customers disagree with their evaluation of the underlying product or service. There are few empirical studies concerning the effects of the variance of online ratings on economic outcomes, and these studies have delivered inconclusive evidence. Whereas Clemons et al. (2006) and Lu et al. (2014) found a positive

---

[1] This work was partially supported by the German Research Foundation (DFG) within the Collaborative Research Centre "On-the-Fly Computing" (SFB 901). I thank Robin Wulfes for excellent student assistance.

[2] Online ratings on Amazon are given on a scale from 1 to 5 stars. For each product, one can observe the number of individual ratings given, the average score of the individual ratings, the distribution of the frequency of the individual ratings (in the form of a bar chart and percentage numbers), and texts that come with each review.

association between the variance of online ratings and sales, Chintagunta et al. (2010) and Ye et al. (2011) did not find support for such an association; in fact, Ye et al. (2009) even found a negative relationship between the variance and sales.

This lack of knowledge represents a major handicap for sellers and buyers alike. Sellers need to understand the implications of the online rating variance so they can adapt the sales strategies of their products and services. Several studies have highlighted that increasing average ratings enables sellers to increase their sales prices (Lewis and Zervas 2016; Neumann and Gutt 2017; Teubner et al. 2017), but it is unclear how sellers should react to changes in the variance of online ratings. Without knowledge of the effects of the variance of online ratings on strategic variables such as price and demand, managers cannot be sure if they should change their prices or alter their demand forecast. Interpreting the variance of online ratings also poses a challenge for buyers because they might not be knowledgeable about whether customer disagreement with a product is a good sign or a bad one. Even though disagreement can be beneficial because negative ratings often contain helpful information (Sen and Lerman 2007), substantial disagreement might scare customers who are pleased by buying mainstream products (Sun 2012).

This study, therefore, aims at empirically analyzing the relationship between the variance of online ratings and sales by testing propositions from an analytical model (Zimmermann et al. 2017). In this model, the online rating variance is divided into different sources—namely, a taste-related share of the variance and a product failure–related share. This study argues that the taste-related variance share can be beneficial for sales, but the failure-related share can be detrimental to sales. This split hinges on the idea that taste-related variance facilitates future customers' purchase decisions; they can find a product that perfectly matches their taste. Taste-related variance, for instance, can be driven by customer disagreement on the ease of navigating the menu of a digital camera—some might like a simple menu, whereas others might prefer a complex menu with highly adjustable settings. However, customers do agree on their dislike of the product failure component of the rating variance; they do not want their digital camera to malfunction. Thus, we pose the following research question:

Does the source of the variance influence the relationship between the variance of online ratings and sales?

To answer our research question, we empirically tested two hypotheses that we adapted from Zimmermann et al. (2017) on a rich data set of 29,332 single online reviews of 1,146 digital cameras from Amazon.com. We employed and validated an unsupervised machine learning approach (Latent Dirichlet Allocation [LDA]) to identify online ratings mentioning product failure in their review texts. This enabled us to calculate the share of the variance related to product failure. We then tested the relationship between this share of failure-related online rating variance and the price as well as sales of the respective product. Our empirical results suggest that there is a significant negative relationship between the failure-related share of the online rating variance with prices and sales of digital cameras. However, prices and sales are positively associated with the taste-related variance share.

Thus, our research makes various substantial contributions to the literature that are accompanied by valuable practical implications to sellers and buyers in online rating communities. First, and to the best of our knowledge, we are the first to empirically disentangle the online rating variance into a taste- and a failure-related share and test the relationship of both shares with sales and prices. In this way, we contribute to the growing literature on online ratings. We provide additional insights into the nature of the relationship between the online rating variance and sales for which studies have presented inconclusive results thus far. Second, our empirical results provide support to the theoretical model of Zimmermann et al. (2017). Therefore, we support the explanatory power of this model that can potentially serve as a reconciliation of conflicting findings in terms of the effects of the online rating variance. Third, our results highlight important practical implications for consumers who rely on user-generated social media to support their purchase decisions. Based on our results, consumers should be able to scan online ratings to infer whether the online rating variance of a product is primarily caused by product failure or by taste-related aspects—e.g., by reading the negative reviews. Fourth, sellers can judge the competitive edge of their online rating variance and adjust prices accordingly. In addition, they can incorporate the rating variance into their sales forecasts. Utilizing our presented approach, online rating systems might assist both sellers and consumers by facilitating the identification of failure ratings.

The remainder of our work is organized as follows: Section 2 presents the related literature, Section 3 discusses the theoretical background, Section 4 presents the research setup, and Section 5 discusses and concludes our study.

2 Related Literature

Our study is related to two substreams of the literature on online ratings: the literature (i) on the impact of textual review content, and the literature (ii) on the relationship between the variance of online ratings and economic outcomes.

First, a nascent stream of research has explored different ways of leveraging the information contained in online review texts. Several studies have demonstrated various functions of textual information in online reviews. They can be used for predicting the pricing power of products (Archak et al. 2011), for the inference and surveillance of market structure (Netzer et al. 2012), and for facilitating trade by signaling benevolence and commitment in peer-to-peer markets (Pavlou and Dimoka 2006). Our study contributes to this stream of research by deploying an automated identification of product failures in online review texts and, even more importantly, validating this approach with manual coders.

Second, some studies have already empirically investigated the economic impact of online rating variance in various contexts. Among these studies, one group has empirically investigated the relationship between the online rating variance and demand for products, such as craft beer (Clemons et al. 2006) or movies (Chintagunta et al. 2010; Moon et al. 2010).[3] Another group of studies has looked at this relationship in the domain of services and exclusively focuses on hotel stays (Lu et al. 2014; Ye al. 2009; Ye et al. 2011).Yet, these studies have delivered contradictory and inconsistent results concerning the relationship between product and service demand and the online rating variance.

With regard to the empirical investigation of product sales, one study finds a substantially positive influence of the online rating variance on product sales investigating the highly differentiated U.S. craft beer market (Clemons et al. 2006). The authors conclude that with hyper-differentiated products, the rating variance helps customers find their most preferred product. Yet, two other studies investigate the effect of the online rating variance on offline movie sales (Chintagunta et al. 2010; Moon et al. 2010) and find no such positive impact of the online rating variance on sales.

Within the group of studies investigating services, the results are mixed. One study finds a positive effect of the online rating variance on hotel sales (Lu et al. 2014). The authors reveal that the positive effect of the online rating variance is strongest for hotels with low classical hotel star ratings. Another study finds no significant influence of the online rating variance on hotel sales (Ye et al. 2011), and a third study shows an overall negative effect of the online rating variance on the online sales of hotels (Ye et al. 2009).

However, to the best of our knowledge, no study so far has investigated the relationship between different sources of the online rating variance—in our case, taste-related and failure-related variance—and sales. Also, none of the prior studies has investigated product failures as one potential source of the online rating variance. In this way, we contribute to the existing empirical literature on online ratings and the online rating variance by investigating the ambiguous relationship between different aspects of the variance and product sales of digital cameras. Empirically, we find that the share of failure-related rating variance is negatively associated with sales, and the share of taste-related variance is positively associated with sales.

[3] We recognize that some might argue that movies watched in theatres might also represent, in part, a service. Yet, movies are digital products (Choi et al. 1997); a particular movie is the same regardless of the service (e.g., friendliness of the staff at the movie theatre), and in consistency with the classification by prior literature (Chintagunta et al. 2010; Moon et al. 2010), we refer to movies as products.

3 Theoretical Background

Theoretically, a large share of products traded on online markets can be described by search and experience attributes (Nelson 1970). *Search attributes* can be inferred prior to purchase by inspecting the product information provided by the manufacturer or seller (Shapiro 1983). For digital cameras, search attributes in product information comprise technical features such as the resolution in megapixels, the item dimensions, and the weight. In contrast to search attributes, *experience attributes* can hardly be learned before the purchase; usually it occurs solely after purchase. For example, for digital cameras, it is difficult to assess the ease of navigating the camera's menu, the filters or scene modes available, and the feel of the material prior to purchase. The products (craft beer and movies) and services (hotel stays) that have been investigated by prior literature share the common trait of many of their attributes being experience attributes; examples include the taste of a beer, the enjoyment of a movie, or the service quality at a hotel.

Yet, online rating systems have changed the way potential consumers can infer experience attributes of products. Online ratings enable potential customers to peer-learn from the digitized word of mouth of other customers (Dellarocas 2003). In the course of this, experience attributes can be transformed into attributes that can be searched and evaluated prior to purchase (Zimmermann et al. 2017). Consequently, potential customers can learn about experience attributes of a particular product without using it (Chen and Xie 2008; Hong et al. 2012; Kwark et al. 2014). For instance, by reading online reviews, potential customers can learn about past customers' experiences concerning the ease of navigating a camera's menu. Some customers might like a simple menu, whereas others might prefer a more complex menu with highly adjustable settings. Customers' disagreements resulting from opposing opinions are thus *taste related*. Potential customers are able to learn not only about the different *taste-related* experience attributes associated with digital cameras, but about product failures. Thus, the risk associated with a failure of the focal product becomes assessable based on the online review texts. This experience attribute is thus *failure related*.

Both types of experience attributes—taste related and failure related—are likely to induce additional variance into the online rating variance of a product. Although some reviewers will rate the digital camera with 5 stars because they prefer a rather complicated camera menu with a multitude of setting options and generally like the product, others might dislike digital cameras with complicated menus and give only 2 or even 1 star for the camera, thus inducing further variance—i.e., increasing the extent to which customers disagree—in terms of online ratings. This increases the taste-related variance share in the total online rating variance. Failure-related aspects also induce variance of online ratings of a product because instances in which digital cameras stop working (product failure) can lead customers to a low product rating. This increases the failure-related variance share in the total online rating variance.

The analytical model by Zimmermann et al. (2017) hinges on the different perceptions of both types of customer variances. The key difference between both sources of variance is that all potential customers would agree about their dislike of failure-related variance, regardless of their taste. The taste-related variance share has two effects: casual photographers or elderly people might prefer digital cameras with simple menu navigation, whereas experienced photographers or technically adept customers might dislike cameras having simple menu options.

3.1 Hypothesis Development

In our research environment, users post reviews about their experiences with digital cameras they bought on Amazon.com. Digital cameras exhibit both search (weight, item dimensions, color) and experience (digital filters/scene modes, reaction time, ease of navigation of the camera menu) attributes. It is also possible for digital cameras to exhibit product failure; the autofocus can stop working, buttons can be dysfunctional, or the camera can simply fail to boot.

For retailers, the composition of the total online rating variance of the camera has important implications for the camera's sales strategy. Prior literature has already revealed that changes in certain metrics of online ratings (such as the average rating) influence the sales strategy; for example, they can lead to price increases or decreases (Lewis and Zervas 2016; Neumann and Gutt 2017). All else being equal—holding the total variance and the average rating constant—an increasing share of taste-related aspects in the total variance of a camera's online rating signals high customer disagreement on the camera. More advanced photographers might enjoy the picture quality of a reflex camera, whereas casual photographers might dislike the complexity with regard to the usage of the camera. In response to that, retailers can increase the price of the camera knowing that advanced photographers are willing to pay more than casual users.

Analogous to that, an increasing share of the failure-related aspects in the variance should lead retailers to decrease their prices because customers generally dislike products that malfunction.[4] Thus, as per Zimmermann et al. (2017), we formulate our first hypothesis:

Hypothesis 1: An increase in the share of taste-related (failure-related) online rating variance is associated with an increase (decrease) in the product's price, holding total variance and average rating constant.

Yet, an increasing share of taste-related online variance also affects the demand for a focal digital camera. The net effect of an increasing share of taste-related variance is determined by two opposing effects. On the one side, customers who like the digital camera but whose preference for simplicity outweighs increased price will not buy it. On the other, there are several reasons why an increased share in taste-related variance may increase demand.

First, an increased share of taste related-variance automatically leads to a decreasing share in failure-related variance. Fewer product failures are appreciated by all potential customers, which should, in turn, increase the demand for the product.

Second, due to the increased taste-related variance, customers with a well-matched taste will learn about the product but may not buy it if it has a lower taste-related variance—holding the average rating constant—because they consider it too "mainstream." Moreover, profit-maximizing retailers will not increase the price up to the point where customer loss due to price increases is larger than the attraction of additional customers because of a high taste-related variance.

Third, a higher taste-related variance enables customers to better decide which product to buy. For example, a customer whose taste is matched with a particular product should feel more comfortable with a product that has a larger taste-related variance than with a product having a lower taste-related variance because he is better informed about the product's pluses and minuses. This should give a potential customer a higher confidence in buying the product.

Fourth, a product with a relatively high taste-related variance might be less suspicious of a product compared to a product with a lower taste-related variance—that consists of relatively positive reviews—when both have the same average rating. The absence of negative reviews for the product with the lower taste variance could undermine the trust in the reviewers by the potential customers. Therefore, we formulate our second hypothesis, as per Zimmerman et al. (2017):

Hypothesis 2: An increase in the share of taste-related (failure-related) online rating variance is associated with an increase (decrease) in the demand for the product, holding total variance and average rating constant.

4 Research Setup

Based on our theoretical background, we hypothesize a positive relationship between the taste-related variance share and prices as well as sales. Equivalently, we hypothesize a negative relationship between the failure-related share of the rating variance and prices and sales. In this section, we empirically test these hypotheses using data on online ratings for digital cameras from Amazon.com.

[4] It is important to note that, when assuming a constant total variance, an increasing share in one aspect of the variance (e.g., taste-related aspects) automatically leads to a decrease in the other aspect of the variance (e.g., failure-related aspects).

4.1 Data

We obtained a data set from Amazon.com containing 29,332 single online reviews of 1,146 digital cameras (McAuley et al. 2015a; McAuley et al. 2015). The online reviews were collected in July 2014 and contain all reviews from May 1999 until the time of collection. Based on the online reviews, we computed the average ratings, the number of ratings, the variance of ratings, the length of the review texts, and the number of helpfulness votes a camera's reviews received. As depicted in Table 1, the digital cameras in our data set have, on average, a rating of 4.067 (AVG_RATING), 25.595 reviews (NUM_REVIEWS), an online rating variance of 1.210 (TOTALVAR), a total of 315 helpfulness votes for their reviews, 198 words per review (AVG_LENGTH), a price of $141.71, and a sales rank of approximately 10,930.

Variable	N	Mean	Std. Dev.	Min	Max
SALES_RANK	1,128	10,930.72	10,086.21	12	155,504
PRICE (in US Dollars)	1,051	141.71	169.489	0.01	899
AVG_RATING	1,146	4.067	0.491	1	5
NUM_REVIEWS	1,146	25.595	33.774	5	280
TOTALVAR	1,146	1.210	0.683	0	3.551
HELPFUL_VOTES	1,146	315	486.134	1	6085
AVG_LENGTH	1,146	198.005	99.332	25.4	937.333

Table 1. Descriptive Statistics

4.2 Identification of Product Failure Ratings[5]

One crucial task in our study is to identify failure ratings in our large data set of online reviews. To this end, we employ probabilistic topic modeling based on LDA. LDA is a widely used unsupervised machine learning method that can identify topics in large collections of documents—in our case online reviews—with written text (Blei 2012; Debortoli et al. 2016). The essential idea behind LDA, according to Blei (2012), is that the authors compose documents *D* by first deciding about a discrete distribution of topics *T* to write about, and then they rely on words *W* from a discrete distribution of words that are typical for the chosen topic. Put differently, a document is defined by a probability distribution over a fixed set of topics, and each topic is defined by a probability distribution over a limited set of words (Debortoli et al. 2016). For each topic of the fixed set of topics, the LDA assigns a probability between 0 and 1 to each document (in this case: online review), indicating how likely it is that this particular document belongs to a certain topic.

The LDA approach has several advantages over alternative approaches that identify topics in written text. First, this approach can handle numerous documents in a very short time. Prior literature, such as in the field of marketing, has traditionally relied on manual coders to identify topics in online reviews (Sridhar and Srinivasan 2012). This approach is time consuming, costly, and difficult to replicate. Our approach circumvents these limitations. Second, it seems plausible that our underlying data suits the LDA assumption that there is a fixed set of topics underlying the documents. Accordingly, recent studies have highlighted the suitability of LDA to analyze online reviews (Debortoli et al. 2016).

Before running the LDA using the web service minemytext.com, we applied standard measures of data preprocessing as suggested in the literature (Debortoli et al. 2016). In particular, we applied stemming to reduce the words to their stem, we used a total of 316 standard stop words (including *the*, *now*, *of*,

[5] A related study by Gutt (2018) draws on the same research environment. Despite overlap in the data, the LDA analysis and the validation, the related study differs in scale and scope, addressing an independent research question.

and *and*) to eliminate common and uninformative words, and we set the n-gram to 1. Because the LDA relies on a fixed number of topics that have to be determined by the researcher before running the analysis, the results obtained can depend on the number of topics chosen. As suggested in the literature (Debortoli et al. 2016), we tried several specifications with between 15 and 80 topics. We evaluated the quality of the results by reading samples of the reviews marked as failure ratings and comparing the mean rating of the failure ratings with the remaining ones. The LDA that yielded the best results after visual inspection comprises 40 topics, which is within the range of 10 and 50 topics usually proposed in the literature (Debortoli et al. 2016). Table 2 provides examples that have been classified as failure ratings by our LDA approach as well as the respective rating given by the reviewer.

Text	Rating
"Shutter button falls off and has a focus lock problem that should have resulted in a product recall. Sony then charges $120 for a repair."	1
"...bought this for when I moved out of the country for a year. The price was good for something that was supposed to tide me over while abroad. It worked, until it stopped working... The lens stopped protruding when the camera was turned on. I do not believe it was anything I did, as I did not drop it or toss it around. I kept it in a camera case. I know it must have been an uncommon defect or event, but I was very disappointed - and cameraless."	2
"I had high hopes for this camera when it arrived today, unfortunately for me the one that I ordered is crazy! The first few times I pressed the power button it would not turn on. After finally getting it to come on I tried taking pictures with it, during the first shot the camera cut off and I couldnt get it to come back on. The camera comes with 2AA duracell ultra batteries and the display showed that it was full when I got it to come back on. I wished I hadnt desired to save money since at the time there was another camera on sale for about twenty more bucks. I had found another of these when searching similar items that is about the same price but after checking targets (the seller) website for reviews I found one in which a reviewer had the exact same problem I am having and theirs ended up with the lens stuck out and useless after only one week, until I read that I had considered giving this one until my birthday at the beginning of the year to act right before returning it.[...]."	2

Table 2. Selected Reviews Classified as Failure Ratings by the LDA

Within the LDA specification of 40 topics, exactly one topic captured the failure ratings, which we identified by reading the most frequently occurring words. Those words were *camera* (7.19%), *len(se)* (3.02%), *problem* (2.77%), *repair* (2.3%), *replace* (1.35%), *fix* (1.3%), and *defect* (0.51%). Because each review has a probability between 0 and 1 for each topic, we identified failure ratings as those reviews that had the highest probability, among all possible topics, for the above-mentioned topic (failure). In total, our LDA analysis identifies 650 failure ratings within the entire data set of 29,332. As expected, the mean rating of the failure ratings (2.13) is substantially lower than the mean rating of the reviews excluding failure ratings (4.24). Figure 1 shows the distribution of failure and nonfailure ratings.

As depicted in Figure 1, the shape of the nonfailure rating distribution resembles the classical J-shape of online reviews (Hu et al. 2017) even though there are fewer 1-star ratings than expected, probably due to filtering out the failure ratings. The distribution of the failure ratings resembles a perfect inverted J-shape. Therefore, failure ratings are associated with substantially lower ratings than nonfailure ratings, probably due to the bad quality and the negative consumption experiences of the customers. However, some 4- and 5-star ratings might occur if product failures happen after a longer period of consumption that satisfy the customer.

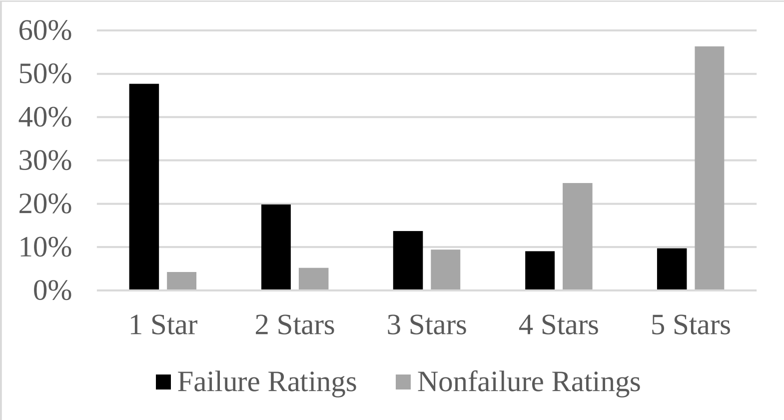

Figure 1. Distribution of Failure and Nonfailure Ratings

4.3 Validation of Identification

In a next step, we would like to validate the results of our topic modeling LDA approach. Our goal is to rule out concerns that our algorithm, for instance, fails to capture some failure ratings or mistakenly classifies ratings as failures that do not mention a product failure. Also, we would like to rule out that irony, sarcasm, or jargon causes our algorithm to misclassify a rating as a failure rating, even though the reviewer was not talking about a product failure. An effective way to do this is by human manual coding and calculating the interrater reliability between the human coding results and the LDA results. Here, we relied on manual human coding conducted independently by two student assistants. The assistants were asked to read a sample of reviews (without other information such as the rating or the brand) and determine whether the reviews mentioned product failures. Because the data contained too many reviews to be handled by humans, we took a random sample of our data comprising 300 reviews (following Lombard et al. 2002). Parametric and nonparametric tests of variables such as SALES_RANK, AVG_RATING, and PRICE did not show statistically significant differences between the entire data set and the random sample; thus, we concluded that the random sample was representative of the whole sample. The results of the agreement between the LDA and the human coders are depicted in Table 3.

	%-Agreement	**N**	**Krippendorff's Alpha**	**Cohen's Kappa**
LDA & C 1	97.99%	300	0.740	0.740
LDA & C 2	98.33%	300	0.806	0.806
LDA & C1 & C2	99.00%	300	0.808	0.875

Table 3. Interrater Agreement Between LDA and Human Coders

At first glance, one can see that the agreement between the LDA and coder 1 (C1) and coder 2 (C2) is high (between 97.99% and 99%). This is reconfirmed by Krippendorff's Alpha and Cohen's Kappa, two conservative standard indices for evaluating interrater agreement (Lombard et al. 2002). Levels of 0.7 for K's Alpha and C's Kappa are sufficient to conclude interrater agreement, and levels above 0.8 indicate large interrater agreement (Lombard et al. 2002). Based on the results of interrater agreement between the human coders and the LDA, we concluded that our automatic topic modeling approach had reliably captured almost all failure ratings contained in our data set. This also reassured us about our choice of the number of topics (40) and the data preprocessing steps; these choices have obviously enabled a robust identification of failure ratings.

4.4 Main Variables

Finally, after the identification of failure ratings, we could compute the percentage share of failure ratings (FAILVARPERC) in the total online rating variance. As shown in equation (1), we calculated

FAILVARPERC as follows. First, we calculated the taste-related variance (TASTEVAR). Then we divided the TASTEVAR by the total variance (TOTALVAR) to obtain the share of taste-related variance in the total variance. We computed TASTEVAR as the squared standard deviation from the mean of the nonfailure ratings, and we computed TOTALVAR as the squared standard deviation from the mean of all ratings (Pindyck and Rubinfeld 2005). Because the remaining share of the variance had to be, by our definition, failure related, we subtracted the taste-related share of the online rating variance from 1 to obtain the failure-related share. We multiplied the result by 100 to obtain a percentage value between 0 and 100.

As an empirical measure for sales, we used Amazon's sales rank. In line with established literature (Chevalier and Mayzlin 2006; Brynjolfsson et al. 2011; Li and Hitt 2008; Meiseberg 2016; Joshi and Mao 2012), we used sales ranks employing the Schnapp – Allwine methodology (LN_SALES_RANK=9.61 – 0.78*ln(SALES_RANK)) to translate the sales rank into the natural log of the sales rank as well-suited proxy for demand. For this measure, higher values were associated with higher sales. Our third variable of interest was the natural logarithm of a camera's price (LN_PRICE).

$$FAILVARPERC = 100 * \left(1 - \frac{TASTEVAR}{TOTALVAR}\right) \tag{1}$$

4.5 Empirical Model

To investigate the relationship between the failure-related variance and the price as well as the demand of a digital camera, we can estimate the following standard linear regression model:

$$Y_i = \beta_0 + \beta_1 FAILVARPERC_i + \gamma X_i + \delta_j + \epsilon_i \tag{2}$$

where the dependent variable Y_i represents the natural logarithm of PRICE, respectively LN_SALES_RANK, of camera i. X_i represents a range of control variables[6], and ϵ_i is the random error term. The key variable of interest in equation (2) is the variable FAILVARPERC, which indicates the share of the failure-related online rating variance of the total online rating variance. Therefore, this variable can take values between 0 and 100, and the coefficient β_1 captures the magnitude of the relationship between the failure-related variance share and price—respectively, sales. We apply the standard ordinary least squares (OLS) assumptions and revisit the distribution of our dependent variables in the robustness checks.

One might argue that, in the case of a digital camera, a major factor that determines the quality of a camera, and therefore the likelihood that a camera will have a product failure, is the camera brand. Naturally, it is possible that a high-class Leica camera will be less prone to product failure than a camera whose brand is not as well established. To mitigate this confounding factor when estimating the relationship between the price, the demand, and the failure-related variance share, we implemented a set of dummy variables δ_j representing brand fixed effects (FE). Brand FE control for all time-constant unobservable heterogeneity—quality differences—between cameras on a brand level.[7] Still, there might be unobservable quality differences in the error term.

4.6 Empirical Analysis

Table 4 displays the regression estimates for our regression model displayed in equation (2). Column (1) presents the results of the model with LN_PRICE as the dependent variable, and column (2) presents the results for the model with LN_SALES_RANK as the dependent variable. Thus, the coefficient β_1 of FAILVARPERC in model (1) tests Hypothesis 1, and β_1 in model (2) tests Hypothesis 2.

---

[6] The control variables are AVG_RATING, NUM_REVIEWS, TOTALVAR, HELPFUL_VOTES, LENGTH, and (when LN_SALES_RANK is the dependent variable) PRICE.

[7] We could not obtain brand information for each camera in our sample. However, our regression results remain qualitatively unchanged when we drop brand FE and run the regression for the whole sample of cameras.

Model	(1)	(2)
VARIABLES	LN_PRICE	LN_SALES_RANK
FAILVARPERC	**–0.00917***	**–0.00415***
	(0.00348)	(0.0015)
TOTALVAR	0.000683	0.00469**
	(0.00271)	(0.00229)
PRICE	–	0.000182
	–	(0.000231)
AVG_RATING	0.249	0.287***
	(0.199)	(0.0767)
NUM_REVIEWS	0.00688	0.00886***
	(0.00419)	(0.00305)
HELPFUL_VOTES	0.00003	–0.000268*
	(0.000204)	(0.000144)
AVG_LENGTH	0.000998	–0.00028
	(0.000923)	(0.000601)
Brand-Level FE	✓	✓
Constant	2.615***	–7.802***
	(0.929)	(0.374)
N	467	460
R^2	0.276	0.418
Note: *** p < 0.01; ** p < 0.05; * p < 0.1.		

Table 4. Main Results

First, the results of column (1) show that the coefficient for FAILVARPERC is negative and statistically significant. The magnitude of the coefficient (β_1: –0.00917, s.e. 0.00348) suggests that the share of failure-related variance is negatively associated with the price of a digital camera. Therefore, we find support for Hypothesis 1. It is important to note that we hypothesize the relationship of both variance components—the taste related and failure related—with the dependent variable. However, both relationships are measured by FAILVARPERC because a 1% increase in the failure-related share of the variance automatically implies a 1% decrease in the taste-related share of the variance. In other words, a 1% increase in the taste-related variance share is +0.00917. Also, unsurprisingly, the association between the total variance and the price is positive. This implies, keeping the share of failure-related variance constant, that an increase in the total variance is associated with an increase in price (because when keeping the failure-share constant and increasing the total variance, the variance must increase in the taste-related share).

In particular, the magnitude of the coefficient of FAILVARPERC is substantial. A 1% increase in the failure-related share of the online rating variance of a camera is associated with a decrease in price by 0.9%. For instance, the price of a camera with a 38% failure-related variance share (the 75th percentile of FAILVARPERC) is 27% (~$38)[8] lower than the price of a camera with an 8% failure-related variance share (the 25th percentile).

Second, the results of column (2) show that the coefficient for FAILVARPERC is negative and statistically significant. Thus, the demand for a digital camera is negatively associated with the share of the failure-related variance (β_1: –0.00415, s.e. 0.0015). Therefore, we also find support for Hypothesis

[8] $38 equals 27% multiplied by 1% of the mean price ($1.41).

2. Again, analogous to this, the taste-related variance share is positively associated with demand. Also in this case, the magnitude of the coefficient is pronounced. In particular, a 1% increase in the failure-related share of the online rating variance of a camera is associated with a decrease in demand by 0.4%. For instance, the demand of a camera with a 38% failure-related variance share (the 75th percentile of FAILVARPERC) is 12.45% lower than the demand of a camera with an 8% failure-related variance share (the 25th percentile).

The total variance is also positively associated with sales; keeping the failure share in the variance constant, an increase in the total variance is associated with more sales. As explained earlier, this is because, when keeping the failure-related share constant, the total variance has to increase in the taste-related share, which is positively associated with sales.

4.7 Robustness Checks

We conducted a series of robustness checks to ensure validity of our results.

First, there was a concern that digital cameras in the early 2000s were systematically different from cameras sold in 2014. Thus, in comparing the failure-related variance of relatively old with new cameras, we compared apples with oranges. To alleviate this concern, we estimated two separate models and display the results in columns (1) and (2) of Table 5. In column (1) we included only cameras with reviews "older" than the median review (which is 2011 in our data set), and in column (2) we included only cameras with reviews newer than the median review. We estimated the models displayed in column (1) and (2) for both dependent variables, LN_PRICE and LN_SALES_RANK. The results remain qualitatively unchanged from our baseline results displayed in Table 4. Due to space limitations, we only display the results for the models with LN_SALES_RANK as dependent variable.

Model	(1) Older Than 2011	(2) Newer Than 2011	(3) Without "Quality" Topics	
VARIABLES	LN_SALES_RANK	LN_SALES_RANK	LN_PRICE	LN_SALES_RANK
FAIL-VARPERC	**–0.00372*** **	**–0.00709** **	**–0.0122*** **	**–0.00681*** **
	(0.00119)	(0.00357)	(0.00419)	(0.00215)
Brand-Level FE	✓	✓	✓	✓
Control Variables	✓	✓	✓	✓
Constant	–7.111***	–8.687***	4.014***	–8.149***
	(0.285)	(0.988)	(1.296)	(0.609)
N	321	139	251	246
R^2	0.355	0.426	0.274	0.422
Note: *** p < 0.01; ** p < 0.05; * p < 0.1. The control variables are the same as in Table 4.				

Table 5. Robustness Checks

Second, there was a concern about our implicit assumption that every review that was not a failure review was automatically said to increase the taste-related variance. Even though this classification is theoretically and logically valid, this may not guarantee that nonfailure reviews really contain taste-related information. While, ultimately, classifying the information contained in the nonfailure reviews is an intricate task, we aim at alleviating these concerns by dropping reviews whose topics are about vertical (*quality-related*) rather than horizontal (*taste-related*) features. The key idea is that every reviewer will appreciate higher quality. Thus, the role of taste should be negligible for reviews mentioning the quality aspects of a camera. Consequently, we dropped all the reviews from topics that had the word "quality" among the 7 most frequently occurring words. We did not drop reviews categorized as failure reviews. We re-estimated our baseline model after removing the "quality" reviews, but our results remained qualitatively unchanged. This gives support to our classification of review in just two categories: taste- and failure-related reviews. Moreover, the coefficient of interest in this specification grew compared to our baseline results, suggesting that we have more actual taste-related reviews in this taste-related variance, which is what we aimed for.

Third, one might argue that validity and the precision of our results are influenced by outliers. To ensure that this is not the case, we re-estimated the model (1) and (2) from Table 4, excluding the 1 percentile and the 99 percentile of PRICE, LN_SALES_RANK, and FAILVAR.

Fourth, one might argue that our results are influenced by our choice of log-transformed variables PRICE and SALES_RANK. To rule out these concerns, we again re-estimated both models without log-transforming these two variables. In the robustness checks, our results remain qualitatively unchanged.

5 Discussion

To a certain extent, our results provide tentative explanations to reconcile conflicting findings of past studies in follow-up studies.

First, and most importantly, one potential explanation for the inconclusive findings of prior studies could be the fact that the analyses conducted in the respective studies did not differentiate between a failure-related and a taste-related share in the variance (Ye et al. 2009; Lu et al. 2014; Ye et al. 2011; Clemons et al. 2006; Moon et al. 2010). As outlined in section 2, prior literature has mostly focused on products—such as craft beer and movies—and services—such as hotel stays. Both of them have two things in common: they comprise a wide range of experience attributes (hotels: friendliness of the staff, taste of the food, comfort of the bed; cameras: product failure, ease of use of the software, usability, feel of the product). Also, both of them can fail. But we recognize that the potential for service failure is likely larger than the potential for product failures with craft beer and movies, as documented in the literature (Webster and Sundaram 1998; Smith and Bolton 1998; Hess Jr. et al. 2003), especially for hotels (Proserpio and Zervas 2017). Consequently, product failure can be leveraged to decompose the total variance in these cases. Although it is impossible to judge ex-post if the different sources of the variance that have been neglected by prior studies constitute the driving force behind the conflicting results, this is one candidate explanation that should be explored in future research. Moreover, the fact that studies investigating the product sales of craft beer and movie tickets have not found a negative relationship between the online rating variance and sales might be due to a limited potential for these products to fail.

Second, our findings are in line with Clemons et al. (2006), although we explore more differentiated facets of the rating variance, and they simply utilize the total variance. We also find that the overall variance is positively associated with sales. Even though beer represents an experience good, one can hardly imagine beer to "fail." Beer can expire, but this should be indicated on the best-before date; therefore, this information should be accessible prior to purchase without relying on online ratings. In that sense, product failures for beer are a search attribute.

Third, our results deliver empirical support for prior experimental (He and Bond 2015) and theoretical studies (Zimmermann et al. 2017). He and Bond's (2015) experimental study investigated the effect of the variance for taste-dissimilar (paintings) and taste-similar (niche computer games) products on purchase intention. They found that a high variance is beneficial for the purchase intention for taste-dissimilar products. In other words, for products that some potential customers like and some others dislike because they have different tastes, the variance increases the purchase intention (He and Bond 2015). For taste-similar products, most of the reviewers should have similar tastes (e.g., players of a certain type of video games). If players observe a high variance of online ratings, their purchase intention decreases because they associate this variance with low product quality. Our results also lend empirical support to this study. In our case, the taste-related variance is the extent to which a digital camera is a taste-dissimilar product. By the same token, the failure-related variance represents the extent to which a camera is taste similar; all customers agree that they dislike product failure.

6 Conclusion

Surprisingly few studies have investigated the relationship between the online rating variance and sales, and their results have been inconclusive. This lack of knowledge is a handicap for researchers, managers, and customers alike. Despite some theoretical accounts, little empirical consensus has been reached in the field of user-generated social media content and, in particular, online ratings. Thus, customer decision-making under disagreement has remained, by and large, a black box. Second, managers need to understand the implications of the online rating variance to adapt the sales strategies of their products and services. Without knowledge of the effects of the variance on strategic variables such as price and demand, managers cannot be sure if they should change their prices or alter their demand forecast. Third, although customers might intuitively know how to interpret the online rating variance, interpreting the online rating variance in combination with the average rating has remained difficult.

To close this knowledge gap, our study focuses on empirically breaking down the relationship between two distinct components of the online rating variance: the taste-related share and the failure-related share with prices and sales. We study this relationship in the context of a consumer good— digital cameras— that can actually fail with a relatively large share of experience attributes. We build upon a theoretical model by Zimmermann et al. (2017) to delineate hypotheses that we later test on our data set of online ratings for digital cameras from Amazon.com. Our empirical results highlight that an increase in the product failure-related share in the online rating variance is negatively associated with prices and sales. The explanation for this is that all customers dislike the failure of a product, which lowers sales and the prices a retailer can charge. In contrast to that, an increase in the taste-related share of the online rating variance is positively associated with prices and sales. This type of customer disagreement helps customers find a product that they really like; when past customers give negative ratings about the difficult usability of a camera, advanced users might prefer this camera over another, simpler camera. Consequently, the retailer can charge higher prices and enjoy a higher demand.

Our results have substantial implications for research and practice. To the best of our knowledge, we are the first to empirically decompose the online rating variance of a product into a failure- and a taste-related component to study the ambiguous relationship of the variance with economic outcomes. Thus, our study helps scholars understand the economic implication of the rating variance. With regard to theory, our results lend empirical support to the model of Zimmermann et al. (2017). In addition, our study lends empirical support to the laboratory findings of He and Bond (2015).

Retailers can benefit from our results to better sell their products. They can adjust their price and demand forecasts based on the observed shares of taste-related and failure-related online rating variance of their product. Moreover, manufacturers can learn about the failures of their own products and improve their design based on the failure ratings.

Customers are able to better incorporate the online rating variance into their purchase decisions. For instance, customers might first inspect negative ratings to check whether they are based on product failure or based on taste-related aspects. Review systems should assist this process by facilitating the identification of failure ratings to help customers find the products they like best.

Naturally, this study also comes with limitations that present avenues for future research. Because our study focuses on consumer goods, a natural extension would be to conduct a study breaking down the online rating variance of services by identifying when services fail (Hess Jr. et al. 2003; Proserpio and Zervas 2017). This might help to explain the conflicting findings by studies on the online rating variance. Also, digital cameras represent a consumer good, and it is unclear how our results play out for digital goods such as apps or video games. Digital goods can also fail–due to bugs or crashes–and future research could study whether taste differences are also beneficial to prices and sales of these goods. Even though our robustness checks alleviate concerns over the comparability of reviews from very old and relatively new cameras and over the classification of the variance in merely two categories, further research in this direction is necessary to address this potential limitation. Moreover, more detailed panel data might be necessary to allow for an analysis of dynamic effects of the variance over time and to identify a more detailed cause-effect relationship.

References

Allen, R. (2016). *Incredible New Stats Show Just How Important Reviews and Ratings Are for Ecommerce Sites*, URL: https://www.smartinsights.com/ecommerce/customer-reviews-social-proof/incredible-new-stats-show-just-crucial-reviews-ecommerce-sites/ (visited 04/13/2018).

Anderson, M., and J. Magruder (2012). "Learning from the Crowd: Regression Discontinuity Estimates of the Effects of an Online Review Database," *The Economic Journal* 122 (563), 957–989.

Archak, N., Ghose, A. and P. G. Ipeirotis (2011). "Deriving the Pricing Power of Product Features by Mining Consumer Reviews," *Management Science* 57 (8), 1485–1509.

Blei D. M. (2012). "Probabilistic Topic Models," *Communications of the ACM* 55 (4), 77–84.

Brynjolfsson, E., Hu, Y. and D. Simester (2011). "Goodbye Pareto Principle, Hello Long Tail: The Effect of Search Costs on the Concentration of Product Sales," *Management Science* 57 (8), 1373–1386.

Cabral, L. and A. Hortaçsu (2010). "The Dynamics of Seller Reputation: Evidence from eBay," *The Journal of Industrial Economics* 58 (1), 54–78.

Chen, Y. and J. Xie (2008). "Online Consumer Review: Word-of-Mouth as a New Element of Marketing Communication Mix," *Management Science* 54 (3), 477–491.

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

Choi, S. Y., Stahl, D. O. and A. B. Whinston (1997). *The Economics of Electronic Commerce*. Indianapolis, IN: Macmillan Technical Publishing.

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

Clemons, E. K., Gao, G. and L. M. Hitt (2006). "When Online Reviews Meet Hyperdifferentiation: A Study of the Craft Beer Industry," *Journal of Management Information Systems* 23 (2), 149–171.

Debortoli, S., Junglas, I., Müller, O. and J. Vom Brocke (2016). "Text Mining for Information Systems Researchers: An Annotated Topic Modeling Tutorial," *Communications of the Association for Information Systems* 39 (7), 110—135.

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

Drewnicki N. (2013). *Survey: 90% Say Positive Reviews Impact Purchase Decisions*, URL: https://www.reviewpro.com/blog/survey-zendesk-mashable-dimensional-research-90-say-positive-reviews-impact-purchase-decisions/ (visited 11/24/2017).

Duan, W., Gu, B. and A. B. Whinston (2008). "Do Online Reviews Matter?—An Empirical Investigation of Panel Data," *Decision Support Systems* 45 (4), 1007–1016.

Gutt, D. (2018). "Sorting Out the Lemons–Identifying Product Failures in Online Reviews and their Relationship with Sales." *Proceedings of the Multikonferenz Wirtschaftsinformatik (MKWI), Lüneburg, Germany*, pp. 1649–1655.

He, S. X. and S. D. Bond (2015). "Why Is the Crowd Divided?: Attribution for Dispersion in Online Word of Mouth," *Journal of Consumer Research* 41 (6), 1509–1527.

Hess Jr., R. L., Ganesan, S. and N. M. Klein (2003). "Service Failure and Recovery: The Impact of Relationship Factors on Customer Satisfaction," *Journal of the Academy of Marketing Science* 31 (2), 127–145.

Hong, Y., Chen, P.-Y. and L. M. Hitt (2012). "Measuring Product Type with Dynamics of Online Product Review Variance," *Proceedings of the 33rd Conference on Information Systems, Orlando, USA*, pp. 2034–2051.

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

Joshi, A. and H. Mao (2012). "Adapting to Succeed?: Leveraging the Brand Equity of Best Sellers to Succeed at the Box Office," *Journal of the Academy of Marketing Science* 40 (4), 558–571.

Koh, N. S., Hu, N. and E. K. Clemons (2010). "Do Online Reviews Reflect a Product's True Perceived Quality? An Investigation of Online Movie Reviews Across Cultures," *Electronic Commerce Research and Applications* 9 (5), 374–385.

Kwark, Y., Chen, J. and S. Raghunathan (2014). "Online Product Reviews: Implications for Retailers and Competing Manufacturers," *Information Systems Research* 25 (1), 93–110.

Lewis, G. and G. Zervas (2016). "The Welfare Impact of Consumer Reviews: A Case Study of the Hotel Industry," *Working Paper*. URL: http://faculty.chicagobooth.edu/workshops/marketing/pdf/tawelfare.pdf (visited 11/27/2017).

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

Lombard, M., Snyder-Duch, J. and C. C. Bracken (2002). "Content Analysis in Mass Communication: Assessment and Reporting of Intercoder Reliability," *Human Communication Research* 28 (4), 587–604.

Lu, Q., Qiang, Y. and R. Law (2014). "Moderating Effects of Product Heterogeneity Between Online Word-of-Mouth and Hotel Sales," *Journal of Electronic Ecommerce Research* 15 (1), 1–12.

Luca M. (2016). "Reviews, Reputation, and Revenue: The Case of Yelp.Com," URL: https://papers.ssrn.com/sol3/Papers.cfm?abstract_id=1928601 (visited 11/27/2017).

McAuley, J., Pandey, R. and J. Leskovec (2015a). "Inferring Networks of Substitutable and Complementary Products," *Proceedings of the 21th In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794.

McAuley, J., Targett, C., Shi, Q. and A. van den Hengel (2015). "Image-Based Recommendations on Styles and Substitutes," *In: Proceedings of the 38th ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 43–52.

Meiseberg B. (2016). "The Effectiveness of E-tailers' Communication Practices in Stimulating Sales of Niche Versus Popular Products," *Journal of Retailing* 92 (3), 319–332.

Moon, S., Bergey, P. K. and D. Iacobucci (2010). "Dynamic Effects Among Movie Ratings, Movie Revenues, and Viewer Satisfaction," *Journal of Marketing* 74 (1), 108–121.

Nelson P. (1970). "Information and Consumer Behavior," *Journal of Political Economy* 78 (2), 311–329.

Netzer, O., Feldman, R., Goldenberg, J. and M. Fresko (2012). "Mine Your Own Business: Market-Structure Surveillance Through Text Mining," *Marketing Science* 31 (3), 521–543.

Neumann, J. and D. Gutt (2017). "A Homeowner's Guide to Airbnb: Theory and Empirical Evidence for Optimal Pricing Conditional on Online Ratings," *Proceedings of the 25th European Conference on Information Systems (ECIS), Guimaraes, Portugal*, pp. 997–1010.

Pavlou, P. A. and A. Dimoka (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), 392–414.

Pindyck, R. and D. Rubinfeld (2005). *Microeconomics*, Upper Sadde River, NJ: Prentice Hall.

Proserpio, D. and G. Zervas (2017). "Online Reputation Management: Estimating the Impact of Management Responses on Consumer Reviews," *Marketing Science* 36 (5), 645–665.

Sen, S. and D. Lerman (2007). "Why Are You Telling Me This? An Examination into Negative Consumer Reviews on the Web," *Journal of Interactive Marketing* 21 (4), 76–94.

Shapiro C. (1983). "Optimal Pricing of Experience Goods," *The Bell Journal of Economics* 14 (2), 497–507.

Smith, A. K. and R. N. Bolton (1998). "An Experimental Investigation of Customer Reactions to Service Failure and Recovery Encounters: Paradox or Peril?" *Journal of Service Research*, 1 (1), 65-81.

Sridhar, S. and R. Srinivasan (2012). "Social Influence Effects in Online Product Ratings," *Journal of Marketing* 76 (5), 70–88.

Sun M. (2012). "How Does the Variance of Product Ratings Matter?" *Management Science* 58 (4), 696–707.

Teubner, T., Hawlitschek, F. and D. Dann (2017). "Price Determinants on Airbnb: How Reputation Pays Off in the Sharing Economy," *Journal of Self-Governance and Management Economics* 5 (4), p. 53—80.

Webster, C. and D. S. Sundaram (1998). Service Consumption Criticality in Failure Recovery. *Journal of Business Research* 41 (2), 153—159.

Ye, Q., Law, R. and B. Gu (2009). "The Impact of Online User Reviews on Hotel Room Sales," *International Journal of Hospitality Management* 28 (1), 180–182.

Ye, Q., Law, R., Gu, B. and W. Chen (2011). "The Influence of User-Generated Content on Traveler Behavior: An Empirical Investigation on the Effects of E-Word-of-Mouth to Hotel Online Bookings," *Computers in Human Behavior* 2 (2), 634–639.

Zimmermann, S., Herrmann, P., Kundisch, D. and B. Nault (2017). "Decomposing the Variance of Consumer Ratings and the Impact on Price and Demand," *Information Systems Research* (Forthcoming). URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3076362 (visited 11/24/2017).