Bargain Hunting on Black Friday – Making Great Deals and Bragging About Them

Kilian Züllig
Ulm University
kilian.zuellig@uni-ulm.de

Stefanie Erlebach
Ulm University
stefanie.erlebach@uni-ulm.de

Alexander Kupfer
University of Innsbruck
alexander.kupfer@uibk.ac.at

Steffen Zimmermann
Ulm University
steffen.zimmermann@uni-ulm.de

Abstract

Online customer reviews (OCRs) are helpful when they provide an unbiased view on a product. Large-scale shopping events (e.g., Black Friday) generate large volumes of OCRs. We hypothesize that OCRs from such events are biased due to price discounts and smart shopper feelings. To test our hypotheses, we analyze OCR data of a large US electronics retailer that emerge from Black Friday purchases and regular purchases. We find that numerical ratings from Black Friday purchases are considerably higher. This effect is also observable in an increase of the average numerical rating through Black Friday purchases. We further observe that textual OCR content from Black Friday purchases focuses more on the purchase conditions (e.g., price discounts) at the expense of other, potentially more helpful content. We further provide managerial implications on how retailers may counteract the negative consequences of such biased OCRs on the quality of their OCR systems.

Keywords: Online Customer Review, Large-Scale Shopping Event, Black Friday, Smart Shopper Feelings

1. Introduction

Events like Black Friday or the Singles Day in China have greatly gained popularity in the last years and significantly increased online retail sales (Essling, 2020). For Black Friday offers in 2021, for instance, Amazon announced to have "more deals than ever before" (Amazon.com, 2021). Each year, TV channels, tech websites and forums greatly discuss the available deals from popular e-commerce retailers and some e-commerce retailers even provide shoppable livestreams on Black Friday (e.g., Amazon Live shopping). All this creates an environment that promotes shopping fever among customers, making Black Friday a large-scale shopping event rather than just a normal price promotion. In the aftermath of such large-scale shopping events, an enormous amount of Online Customer Reviews (OCRs) is written as a simple consequence of the high sales during the event. As anecdotal evidence from our dataset, around three times as many OCRs were published in the month following Black Friday in 2020 compared to the average months before on the Best Buy website.

For OCR systems, it would be critical if the OCRs that originate from purchases during large-scale shopping events are biased as the mere quantity of generated OCRs could affect aggregated OCR metrics such as the average rating of products. Since previous research already observed that regular price promotions such as coupons influence OCRs (e.g., Wu et al., 2021; Zhu et al., 2019), these concerns are valid. In particular, there are reasons to assume that these events feature specific economic and emotional components for customers that are likely to be represented in the resulting OCRs as well. Statements in OCRs like, for instance, "I am super happy with the purchase! I got the product at a great price during Black Friday." focus more on aspects of the purchase conditions than on the product itself. In this example, the positive textual content of the OCR is driven by the large-scale shopping event but not by product quality. If such OCRs frequently occur after Black Friday, the effect on the OCR system is substantial. In other words, large-scale shopping events might substantially bias OCRs of the offered products and potentially damage the helpfulness of OCR systems.

In this paper, we therefore examine the impact of a large-scale shopping event on OCRs. Thereby, we complement existing research on the effect of regular price promotions on OCRs by additionally considering potential emotional aspects during large-scale shopping

URI: https://hdl.handle.net/10125/103114
978-0-9981331-6-4

events. In particular, we aim to address the following research question:

RQ: How do large-scale shopping events affect the numerical rating and textual content of OCRs?

We develop our hypotheses along the economic components of large-scale shopping events (i.e., the price discount) as well as their emotional component (i.e., happiness and pride for having snatched a deal). For the economic component, we hypothesize higher satisfaction due to the lower prices paid by using the concept of perceived customer value (see, e.g., Sánchez-Fernández & Iniesta-Bonillo, 2007). For the emotional component, we draw on the appraisal theory (Frijda, 1987; Lazarus, 1991; Roseman, 1991) and hypothesize that so-called smart shopper feelings (Schindler, 1998) arise during large-scale shopping events. Although both components are expected to increase the numerical rating, we further expect the emotional component of large-scale shopping events to also impact the textual content of OCRs.

To examine our hypotheses, we collected a large dataset of OCRs from the US electronics retailer Best Buy. The dataset includes OCRs from Black Friday purchases (i.e., purchases made during the Black Friday and Cyber Monday shopping event in 2020) and, for comparison, also OCRs that originate from regular purchases (i.e., purchases not made during Black Friday). By matching our dataset at the product level, we can identify effects on OCRs that arise during large-scale shopping events. We examine the numerical rating of OCRs as well as the textual content of OCRs. For the latter, we applied a supervised topic model to learn about the composition of different aspects in the textual content of OCRs.

The results of our study indicate that OCRs are significantly affected by large-scale shopping events: Numerical ratings from purchases during large-scale shopping events are significantly higher compared to regular purchases. This effect is particularly strong for products with a low average numerical rating. We also find that the sum of the many individual numerical ratings that are influenced by Black Friday discounts lead to a significant increase of the overall average ratings of the promoted products. This highlights the relevance of large-scale shopping events as source of bias in OCR systems. Furthermore, large-scale shopping events also have an impact on the textual content of OCRs. More specifically, we observe that OCRs emerging from large-scale shopping events address more purchase aspects and that the overall proportion of purchase aspects increases as well. An additional analysis further suggests that this increase of content about the purchase comes at the expense of product-, customer-, and seller-related content in OCRs.

Our study contributes to existing research by providing new insights on how large-scale shopping events influence OCRs. This has important implications for both, theory and practitioners. From a theoretical perspective, our findings reveal that an emotional component (i.e., smart shopper feelings) of large-scale shopping events impacts OCRs, in addition to the economic component. We thereby extend previous research on the effects of regular price promotions on OCRs by highlighting that smart shopper feelings are present and reflected in OCRs emerging from large-scale shopping events.

For practitioners, our findings suggest that large-scale shopping events can be detrimental for the quality of OCR systems as they inflate numerical ratings and generate potentially less helpful OCRs. Hence, if OCR system providers want to ensure high quality OCRs, they should be aware of the negative consequences of large-scale shopping events and take potential countermeasures, e.g., by means of review templates. On the other hand, policy makers or consumer protection agencies should be aware, that e-commerce retailers can also systematically use large-scale shopping events to quickly generate a high OCR volume for newly launched products and to “improve” the numerical rating of products with mediocre or low quality.

2. Related Literature

The impact of ordinary price promotions (e.g., coupons) on OCRs has mainly been investigated in the context of the hospitality industry, i.e., restaurants or hotels. Byers et al. (2012), for instance, observe that users of coupons post lower numerical ratings for restaurants than regular customers. Li (2016) extends the work of Byers et al. (2012) and finds, based on the same dataset, that coupons improve numerical ratings for restaurants with a low existing average numerical rating and a low OCR volume. Similarly, Zhu et al. (2019) use OCR data of restaurants from an OCR platform and find a positive impact of coupons on numerical ratings independent of the existing average rating. In addition, the authors also examine the effect of coupons on textual content of OCRs and find that guests using coupons write less about their dining experience.

Besides the inconclusive evidence on the effect of promotions on numerical ratings (Byers et al., 2012; Li, 2016), we also expect the effects of coupons in the hospitality industry to differ from the overall effects of promotions of e-commerce retailers: Redeeming a coupon in a restaurant with personal interaction with the staff or the business owner might have different emotional consequences than snatching an offer from a large, anonymous e-commerce retailer.

Studies that investigate the effect of promotions on OCRs within the setting of e-commerce retailers are rare and we are only aware of the study by Wu et al. (2021). The authors examine transactional data from an e-commerce retailer that also includes OCRs from purchases where coupons for a price discount were applied. The authors find a positive impact of coupons on the numerical ratings. They further observe that the monetary savings are addressed in the textual content of OCRs. Although content related to monetary savings is observable in the short-run, it disappears in the long-run.

The findings by Wu et al. (2021), however, cannot be simply transferred to our setting as a coupon that is offered to an individual customer is likely to have different emotional consequences than large-scale shopping events. Hence, there is a lack of knowledge on how large-scale shopping events affect OCRs. Understanding these effects is of particular importance because large-scale shopping events have the potential to significantly influence aggregated metrics such as the average numerical rating. This is to expect from their vast sales volume, which is not comparable with individual coupons. By explicitly considering such an event (i.e., Black Friday), we extend existing research on the effect of price promotions on OCRs. In the next section, we outline why we expect the additional emotional component of large-scale shopping events to be an important determinant for the textual content of OCRs which cannot be observed in regular coupon promotions.

3. Theoretical Background and Hypotheses Development

We expect that the effects of large-scale shopping events on OCRs have both, an economic component and an emotional component. To explain the economic component, we use the concept of perceived customer value (see Sánchez-Fernández & Iniesta-Bonillo, 2007 for a review) and for the emotional component, we draw on appraisal theory (Frijda, 1987; Lazarus, 1991; Roseman, 1991).

Starting with the economic component, perceived customer value is used in marketing literature to explain customer behavior and purchase decisions. It represents a central part of the means-end model by Zeithaml (1988) where "perceived value is the consumers overall assessment of the utility of a product based on perceptions of what is received and what is given" (Zeithaml, 1988, p. 14). It is the trade-off between the perceived quality of a product and the perceived sacrifice which must be made to obtain it (i.e., paying the price). Grewal et al. (1998) extend the model for price discounts and argue that the reduction of a price (compared to the original reference price) is perceived as a lower sacrifice and consequently increases the perceived value of a product, which in turn is a determinant for customer satisfaction (McDougall & Levesque, 2000).

The appraisal theory, on the other hand, explains the formation of emotions and action tendencies based on the evaluation of events (Lazarus, 1991; Frijda, 1987; Roseman, 1991). This theory states that emotions are not the direct result of an event but are formed by the evaluation of the event and its outcome based on appraisals. These appraisals are defined along multiple dimensions, of which goal congruency and attribution of agency are the most relevant in the context of large-scale shopping events. Goal congruency determines the valence of the emotion. An event that is goal congruent (i.e., consistent with the personal goals) causes positive emotions like happiness and joy (Lazarus, 1991). Attribution of agency additionally refers to the perceived responsibility for the event and its outcome. When a positive outcome of an event (e.g., getting a bargain) is attributed to oneself, a positive emotion of pride is created (Weiner, 1985).

Since large-scale shopping events like Black Friday come along with advertisements that prominently highlight deep discounts, limited quantities and limited time of the promotion, a feeling of competition between customers is created (Aggarwal et al., 2011) and bargain hunting is promoted. A customer who successfully hunts a bargain attributes this positive outcome to her own actions, resulting in "the pridelike satisfaction of having won in an implied game against the seller (Rose, 1988) and […] other consumers" (Schindler, 1998, p. 388). Schindler (1998) calls these emotions, which are a consequence from perceiving oneself responsible for the benefits of a discount, smart shopper feelings.

Hence, both the increased perceived value, as the economic component of large-scale shopping events, and the smart shopper feelings, as the additional emotional component, result in high satisfaction and happiness with the purchase. This satisfaction and happiness should in turn be expressed in OCRs when customers assign numerical ratings. As a consequence, we expect Black Friday OCRs to have a higher numerical rating than regular OCRs and state our first hypothesis as follows.

H1: Large-scale shopping events increase the numerical ratings of OCRs.

The positive emotions (i.e., happiness, satisfaction and pride) that are postulated by the appraisal theory further trigger proactive action tendencies (Frijda, 1987). In the context of large-scale shopping events, such proactive action tendencies are reflected in the impulse of a customer to brag to others about her success (Folkes, 1988). In other words, all these emotions (i.e., smart shopper feelings) increase the motivation to share the own positive experiences with the purchase and the shopping event (Gelbrich, 2011; Bicen & Madhavaram, 2013; Schindler, 1998). Hence, we expect that large-scale shopping events influence the textual content of OCRs due to the emotional aspects of smart shopper feelings. In more detail, we expect that the textual content of OCRs which originate from purchases during large-scale shopping events include more aspects related to the purchase itself. Therefore, our second hypothesis reads as follows.

H2: Large-scale shopping events increase textual content of OCRs about purchase aspects.

If large-scale shopping events trigger purchase-related content in OCRs primarily due to the proactive behavior from smart shopper feelings, these purchase aspects in OCRs could be considered as proxy for smart shopper feelings. As these purchase aspects reflect the positive experiences with the purchase during the large-scale shopping event, they should positively influence numerical ratings of OCRs as well. We therefore expect a mediating effect of purchase-related aspects between the large-scale shopping events and the numerical rating that is driven by the emotional component of such events. Hence, we hypothesize the following.

H3: Textual content of OCRs about purchase aspects partially mediates the effects of large-scale shopping events on numerical ratings of OCRs.

4. Research Environment and Dataset

To empirically test our hypotheses, we collected data on products and their OCRs from the e-commerce website of the US electronics retailer Best Buy. All collected products were on sale during the Black Friday and Cyber Monday shopping event in 2020. We focus our analysis on TVs, as they represent a very homogenous product group which facilitates the textual modelling of key features in OCRs (see Section 5 for more details).

For our data collection, we developed a specific web scraper for the Best Buy website in Python. The data collection process was divided in two steps. In step one, we monitored the Best Buy website during the Black Friday promotion period and stored data on all available products from the TV category that were promoted. We gathered product and promotion related variables such as a unique product ID, product name, brand, price, and discount. In step two, we then extracted all existing OCRs including numerical rating (i.e., 1-star to 5-stars), OCR date, OCR title and textual content of all TVs that were collected in step one. As we also gathered the time the product was owned before the OCR was published, we can identify whether a respective OCR emerged from a Black Friday purchase or not. To reduce the risk of biases due to fake OCRs, we only considered OCRs of verified purchases.

With the underlying data, we can define a treatment group and a control group of OCRs that are matched on product level. The treatment group contains Black Friday OCRs, i.e., OCRs that correspond to purchases during Black Friday in 2020. The control group consists of regular OCRs corresponding to regular purchases of the same products from a period before, that does not contain any large-scale shopping event (i.e., between 2020-07-15 and 2020-10-15).[1] This allows us to measure differences in our dependent variables caused by the large-scale shopping event. As we collected OCR data in January 2021, the treatment group by definition includes only OCRs of customers that owned the product for two months or less. Therefore, we also restricted the control group to OCRs that were written within two months of product use to avoid that product evaluation differs depending on usage time.

Our final sample for the analysis is based on 29,327 OCRs from 115 products. The treatment group includes 15,022 OCRs, and the control group consists of 14,305 OCRs.

5. Supervised Modeling of Textual OCR Content

As we hypothesize in H2 that large-scale shopping events increase textual content about purchase aspects in OCRs, we need to capture the composition of different aspects that are mentioned in OCRs. We thereby build upon the framework by Zhu et al. (2017) that identifies three commonly addressed aspects in textual content of OCRs, namely, product (i.e., product quality, aesthetics, functionality, as well as price and promotions), seller (i.e., seller trustworthiness, service quality, and logistics quality) and customer aspects (i.e., emotional attitudes, recommendation expressions, and attitudinal loyalty).

Although we mainly agree on this classification, we argue for separating content related to price or promotions from product aspects: product price can vary strongly between customers, between retailers, and over time (e.g., due to price promotions and large-scale shopping events), it is therefore an attribute of the individual offer and not of the product itself.

[1] The relevant period for the control group was selected in a way that we obtain a similar sample size as for the treatment group and is not too close to Black Friday in November.

Textual content addressing the price and promotions should hence be treated independently from the other product aspects. Consequently, we adapt Zhu et al.'s (2017) framework by adding a fourth aspect that explicitly describes individual purchase conditions. In our understanding, this aspect subsumes all references to product prices, deals, promotions etc. Hence, our revised framework to capture the composition of textual content in OCRs consists of four different aspects (i.e., purchase, product, seller, customer). Table 1 shows our adapted framework.

Because we expect that purchase aspects are more prevalent in OCRs from large-scale shopping events, it is necessary to categorize the textual content as precisely as possible in one of the four aspects. In other words, distinctions between terms like "picture quality" or "service quality" are of great importance. Although unsupervised topic models such as Latent Dirichlet Allocation (Blei et al., 2003) are often applied to examine textual content of OCRs, these models are not specific enough by their unsupervised nature and can struggle with short texts (Cheng et al., 2014). Since we know the definitions of the aspects that we expect to be mentioned in the textual content of OCRs, we can directly apply a supervised topic modelling approach to train a model that can specifically capture our four predefined aspects.

To create a labeled training dataset for the supervised topic model, we proceed as follows: First, because the textual content of one OCR can potentially address multiple aspects, we split all OCRs into single clauses which are more likely to address only one specific aspect. Then, we use 2,000 clauses of randomly sampled OCRs for labeling. Two human coders, who have extensive experiences with online shopping and OCRs, manually assigned each clause to one of the four aspects from our framework (i.e., purchase, product, seller, customer). If a clause does not match any of these aspects, it is assigned to an additional default category "Other". The quality and objectiveness of the labeled data is assessed by percent agreement, a measurement of intercoder-reliability to quantify the extend of agreement between or among coders (Lombard et al., 2002). Our data achieved a percent agreement of 84%, suggesting an acceptable level of agreement among our two coders (Neuendorf, 2002). In cases of disagreement between the two coders, we consulted an additional third human coder and used the majority vote of the three independent labels to determine the final aspect. If all three coders disagree (only applies to 2% of the clauses) the respective clause is assigned to the default category "Other".

With the labeled training dataset, we then apply the supervised topic model. We opt for convolutional neural

Table 1. Definitions of relevant aspects of textual OCR content

Aspect	Components	Definition
Purchase	Purchase conditions	Information regarding the individual purchase conditions such as value, paid price, discounts, promotions, etc.
Product	Product functionality	Descriptions regarding product usage, performance, usefulness, etc.
	Product quality	Descriptions regarding the quality of a product, such as product durability, product conformity, etc.
	Product aesthetics	Descriptions regarding product appearance, such as product package, product design, etc.
Seller	Seller trustworthiness	Descriptions regarding product authenticity and product freshness.
	Logistics quality	Descriptions regarding logistics quality provided by the website, such as delivery speed, delivery accuracy, etc.
	Service quality	Descriptions regarding services provided by the website, such as payment choice, return and refund services.
Customer	Emotional attitudes	Emotional descriptions that express personal or others' feelings.
	Recommendation expressions	Expressions about advising others to buy or not to buy a product.
	Attitudinal loyalty	Expressions regarding predisposition, commitment and attitudinal preference towards a product and the willingness to repurchase it.

Note: All definitions closely follow Zhu et al. (2017). "Product" and "Purchase" aspects are adapted.

networks (CNNs) as they perform particularly well for sentence-level topic predictions and make feature engineering obsolete (Minaee et al., 2021). More specifically, we use a CNN model developed by Kalchbrenner et al. (2014). Before training the model, we perform a spell check and remove special characters from the clauses. More sophisticated text preprocessing is not required by Kalchbrenner et al.'s (2014) model. We then use 90% of the 2,000 manually labeled clauses as training data and keep 10% as out-of-sample test data. Based on this test data, the CNN model achieves 83% accuracy which leads us to conclude that the model performs well in the labeling task. Therefore, we let the trained CNN model classify the remaining unlabeled clauses of the complete dataset.

6. Statistical Analysis and Results

We start our analysis by investigating the overall effect of the large-scale shopping event on the numerical ratings of OCRs (H1). For this purpose, we fit a fixed effects model to the data, which is specified by

$$Rating_{ij} = \beta_1 BlackFriday_{ij} + \gamma_i + \epsilon_{ij}. \quad (1)$$

The dependent variable $Rating_{ij}$ represents the numerical rating of OCR j for product i. As independent variable, we include the dummy $BlackFriday_{ij}$ that codes the treatment and the control group (i.e., being one if an OCR originates from a purchase during the large-scale shopping event and zero otherwise). We implicitly control for product-specific variables by including product-fixed effects γ_i. Finally, ϵ_{ij} represents the remaining error term. In all regression models, we use robust bootstrapping standard errors to account for deviations from the normality assumption of the OLS estimator.

The results are presented in column (i) of Table 2. We observe a significant increase in the numerical rating. While the magnitude of the change seems to be rather small, it is important to keep in mind that the dataset generally has a high average numerical rating (4.68), and ratings are limited by the upper bound of 5. Consequently, there is only little room for improvement. To account for this fact, we allow for variation in the effect strength between products with high and low existing average numerical rating by fitting the model specified by

$$Rating_{ij} = \beta_1 BlackFriday_{ij} + \beta_2 BlackFriday_{ij} \times AvgRating_i + \gamma_i + \epsilon_{ij}. \quad (2)$$

We additionally include the interaction term $BlackFriday_{ij} \times AvgRating_i$. $AvgRating_i$ depicts the existing average numerical rating of product i before Black Friday.

The results for this exercise are shown in column (ii) of Table 2. As expected, we now observe a considerably higher coefficient of $BlackFriday_{ij}$. The negative (and statistically significant) coefficient of the interaction term implies that the effect of large-scale shopping events is stronger for products with a low average numerical rating. In sum, both results in columns (i) and (ii) of Table 2 support H1 and suggest that large-scale shopping events significantly increase numerical ratings of OCRs.

To investigate whether large-scale shopping events increase textual content about purchase aspects in OCRs (H2), we use two measures. First, we use the proportion to which textual content of OCRs consists of purchase aspects. This is calculated by dividing the length of each clause about the purchase aspect by the total length of the textual content of the respective OCR. Second, we use the (relative) frequency of OCRs that include purchase aspects. This is simply the number of OCRs containing clauses that address purchase aspects, divided by the total number of OCRs. For both variables, we examine the differences between the

Table 2. Regression results

	(i)	(ii)	(iii)	(iv)
	$Rating$			
$BlackFriday$	0.028 ** (0.009)	1.413 *** (0.402)		0.025 * (0.010)
$BlackFriday \times AvgRating$		-0.299 *** (0.087)		
$Purchase$			0.202 *** (0.022)	0.201 *** (0.023)
F-Statistic	3.583 ***	3.638 ***	4.528 ***	4.544 ***
Adjusted R^2	0.010	0.010	0.014	0.014

*Note: ***, ** and * indicate statistical significance at the 0.1%, 1% and 5% level, respectively. Robust bootstrapping standard errors are shown in parentheses.*

treatment group and the control group. To obtain more robust results, we excluded OCRs with text length larger than 484 words (99.9% quantile), as they represent extreme outliers that might bias our results.

Based on independent two-sample t-tests, we observe both a significant increase in the proportion of purchase aspects (+2.6%, $p < 0.001$) and in the frequency of purchase aspects (+6.4%, $p < 0.001$) in textual content of OCRs. This implies that customers from large-scale shopping events generate significantly more volume of textual content about purchase aspects relative to the overall volume of textual content and write significantly more frequently about purchase aspects. This evidence supports H2.

To test H3, we perform a mediation analysis as implemented in the PROCESS macro (Model 4) for SPSS (Hayes, 2017). Because the PROCESS macro does not support fixed effects regression models and we use panel data, we replicate the same procedure in STATA for fixed effects regressions.[2] As in the PROCESS macro, we also assess statistical significance of the indirect effect via a bootstrapping procedure. According to Hayes (2017), the bootstrapping confidence interval tends to have greater statistical power compared to the traditional test by Sobel (1982). To start with, we estimate the effect of the hypothesized mediating variable on the numerical rating with

$$Rating_{ij} = \beta_1 Purchase_{ij} + \gamma_i + \epsilon_{ij}. \quad (3)$$

$Purchase_{ij}$ represents the proportion of purchase aspects in textual content of OCRs, as outlined in the analysis of H2 before.[3]

The positive and significant coefficient of 0.202 in column (iii) of Table 2 confirms the expected positive effect of purchase aspects on the numerical rating. We then estimate the direct effect of the large-scale shopping event on the numerical rating with the model specified by

$$Rating_{ij} = \beta_1 BlackFriday_{ij} + Purchase_{ij} + \gamma_i + \epsilon_{ij}, \quad (4)$$

which also includes the dummy $BlackFriday_{ij}$ in addition to $Purchase_{ij}$. Column (iv) of Table 2 indicates a significant direct effect of 0.025 for Black Friday. Hence, the resulting indirect effect yields 0.003 (i.e., the difference between the total effect of 0.028 in column (i) of Table 2 and the direct effect of 0.025 in column (iv) of Table 2). Based on 5,000 bootstrapping samples, we can confirm that the indirect effect is statistically significant ($p < 0.001$). Hence, we can also support H3.

7. Consequences on the Overall Textual Content of OCRs

As we confirm H2 in the previous section (i.e., large-scale shopping events increase textual content about purchase aspects in OCRs), we now want to examine the consequences on the overall textual content of OCRs. Hence, we examine whether the increase of purchase-related aspects is simply due to additional content or if it comes at the expense of other, potentially more helpful aspects.

For this purpose, we not only compare the textual content about purchase aspects in OCRs between the treatment and the control group but also consider the relevant product, seller, and customer aspect. As these three aspects are considered potentially helpful to customers (Sun et al., 2019), we examine them jointly as one aggregated group of helpful content.

Table 3. Change of textual OCR content due to Black Friday

Panel A: Proportion of Aspects			
	Black Friday	Regular Purchase	Difference
Purchase	12.8%	10.2%	+2.6% ***
Product, Customer, Seller	79.5%	81.3%	-1.8% ***
Other	8.5%	7.8%	-0.7% ***
Panel B: Frequency of Aspects			
	Black Friday	Regular Purchase	Difference
Purchase	30.6%	24.1%	+6.5% ***
Product, Customer, Seller	95.2%	95.8%	-0.6% **
Other	27.2%	29.7%	-2.5% ***

*Note: ***, ** and * indicate statistical significance at the 0.1%, 1% and 5% level, based on t-tests (Panel A) and χ^2-tests (Panel B), respectively.*

Table 3 Panel A shows the proportion that the aspects have in the textual contents in OCRs from Black Friday purchases and OCRs from regular purchases. The last column displays the difference and whether the difference is statistically significant based on an

[2] For robustness, we also performed the mediation analysis in SPSS using the PROCESS macro without fixed effects. The results are qualitatively equivalent to those in Table 2.

[3] Conducting the mediation analysis with the second metric (i.e., the frequency of purchase aspects) leads to qualitatively equivalent results (not tabulated).

independent two-sample t-test. The first row shows the tabulated results for H2 above. The second row indicates that the increase of purchase aspects is at the expense of other aspects: The proportion to which textual content of OCRs consists of product, customer or seller aspects is significantly lower for OCRs from Black Friday purchases compared to OCRs from regular purchases. Textual content assigned to the default category "Other" slightly decreases as well.

Table 3 Panel B shows the (relative) frequency in which the aspects are addressed in OCRs from Black Friday purchases and OCRs from regular purchases. The last column displays the difference and whether the difference is statistically significant. Because we compare frequencies, we apply a χ^2-test here. As for Panel A above, the first row shows the tabulated results for H2. The second row again shows that the increase of purchase aspects is at the expense of the other, potentially helpful aspects: The frequency in which OCRs address product, customer or seller aspects is significantly lower for OCRs from Black Friday purchases compared to OCRs from regular purchases.

Finally, we examine whether OCRs that include purchase aspects are significantly different than OCRs that do not include purchase aspects in general. As presented in Table 4, we observe that OCRs which address purchase aspects generally contain significantly less textual content about potentially helpful aspects, measured in the total number of words addressing each aspect. Put differently, OCRs that address purchase aspects include, on average, around four words less about potentially helpful aspects than OCRs that do not address purchase aspects.

Table 4. Average number of words of other aspects depending on purchase aspects

	Purchase included	Purchase not incl.	Difference
Product, Customer, Seller	21.77	25.32	-3.54 ***
Other	3.11	2.99	-0.12

*Note: ***, ** and * indicate statistical significance at the 0.1%, 1% and 5% level, based on t-tests.*

8. Discussion and Conclusion

Our findings suggest that large-scale shopping events influence OCRs in a systematic way. First, by using a dataset that is matched on product level, we observe that customers tend to post OCRs with a higher numerical rating when they purchase the same product during a large-scale shopping event. This is especially true for products that have a low average numerical rating as these products have the potential to increase their rating. Products with high existing average numerical rating, on the other hand, are less affected as their generally high numerical ratings leave little room for improvement. Consequently, the exogenous shock on the numerical rating due to a large-scale shopping event reduces the differences between numerical ratings of different products. This limits the ability of customers to quickly distinguish between high-quality products and low-quality products. To emphasize the relevance of these effects, we examine the average numerical ratings of all available OCRs per product, before and after Black Friday. Although the increase in the average numerical rating is only marginal for the overall sample (i.e., before: 4.635; after: 4.649; $p<0.01$), it is still significant. Similar to the effects on the individual ratings, however, the increase for products with the 10%-lowest average rating is much larger at 0.1 (i.e., before: 4.272; after: 4.374; $p<0.001$). This constitutes a significant bias introduced by the large-scale shopping event.

Second, the textual content of customers' OCRs includes significantly more purchase aspects when they buy a product during a large-scale shopping event. Our additional analysis in Section 7 suggests that this increase of purchase aspects comes at the expense of product, customer, and seller aspects in OCRs which are considered to be particularly helpful (Sun et al., 2019). Hence, OCRs from Black Friday purchases might be less helpful and therefore have the potential to negatively affect the overall quality of an OCR system.

Third, the effect of large-scale shopping events on the numerical rating of OCRs is partly mediated by the increase of purchase aspects in textual content. This supports our theoretical assumptions that, using purchase aspects in textual content of OCRs as a proxy for smart shopper feelings, not only the economic component of large-scale shopping events influences numerical ratings but that large-scale shopping events also trigger an emotional component that is reflected in the numerical ratings as well.

Theoretical Contribution: With this study, we contribute to IS research by providing a deeper understanding about the effects of large-scale shopping events on OCRs. Our findings suggest that the emotional component, which can be explained by the appraisal theory (Frijda, 1987; Lazarus, 1991; Roseman, 1991), contributes to the effects on OCRs. In more detail, we complement existing research on the effect of regular price promotions on OCRs (e.g., Wu et al., 2021; Zhu et al., 2019) by highlighting that smart shopper feelings during large-scale shopping events (in addition to the economic component) positively impact the numerical ratings of OCRs. This clearly distinguishes our results from Wu et al. (2021) as individual coupons are unlikely to elicit bargain hunting and smart shopper feelings. By adapting Zhu et al.'s (2017) framework for the composition of OCRs' textual content, we further observe that the emotional component is also reflected in the textual content of OCRs in form of increased occurrence of purchase aspects.

Managerial Implications: Our results indicate that large-scale shopping events significantly influence OCRs. For an (independent) OCR system provider that aims to provide unbiased OCRs, large-scale shopping events have negative consequences: They generate a large volume of OCRs with numerical ratings that are inflated by the economic and emotional components and with textual content that addresses less of the potentially helpful aspects for customers. Thus, large-scale shopping events seem to be detrimental for the quality of OCR systems as source of information. If an OCR system provider tries to avoid negative consequences of such purchase-focused OCRs, adequate measures should be taken to maintain the quality of the OCR system. For example, to ensure that relevant aspects are mentioned, customer's attention could be deliberately drawn to the purchased product by asking for relevant aspects through the provision of OCR templates (Poniatowski et al., 2019). Providers of OCR systems could further implement design features like explicitly labelling OCRs emerging from large-scale shopping events or excluding them from the calculation of the average numerical rating.

Policy makers or consumer protection agencies should be aware, that e-commerce retailers can, however, also make systematic use of large-scale shopping events: If, for instance, a retailer's goal is to sell products with mediocre or low quality, their reputation can be enhanced by offering special deals during such events. This makes their numerical ratings converge towards the numerical rating of high-quality products and thus potentially result in higher sales in the long run (Chevalier & Mayzlin, 2006). Another side-effect of large-scale shopping events is the fact that a massive number of OCRs is generated during these days. Hence, these events can also be used to quickly generate high OCR volume for newly launched products which might normally take a longer time. This is of economic relevance as well, because the number of existing OCRs represents another important determinant for customers' purchase decisions (Duan et al., 2008).

Limitations and Future Research: Although our study provides several important insights, it does not come without some limitations which might, however, serve as starting points for future research. First, differences in characteristics between product categories might have an impact on the effects of large-scale shopping events on OCRs, which we cannot observe due to our focus on TVs. Future research could build on this point and examine the effects of large-scale shopping events for other product categories as well. Second, our analyses are based on a real-world dataset, and we match OCRs on product level to observe whether customers from large-scale shopping events review differently. Due to our data collection, we can, however, not account for potential price discounts in our control group. Furthermore, applying a standard diff-in-diff approach is hardly possible since most of the products are promoted during Black Friday and thus there is an insufficient number of products to construct a control group. Third, even though purchase aspects in textual content of OCRs can be considered as proxy for smart shopper feelings, it is impossible to perfectly identify smart shopper feelings with a real-world dataset. Future research could address these limitations by conducting a laboratory experiment to properly define a control group and to explicitly ask customers about their smart shopper feelings. Using this information for examining the effect on OCRs could allow to better isolate the impact of the economic and emotional components. Similarly, we acknowledge, that our effect sizes are – although significant – rather small. As this can also be a consequence of a real-world dataset, a laboratory experiment that isolates the effect of smart shopper feelings might result in larger effect sizes.

As we observe that the composition of different aspects in the textual content of OCRs differs between OCRs from Black Friday purchases and OCRs from regular purchases, we conjecture that OCRs that originate from purchases during large-scale shopping events are less helpful. Although this argument is backed by existing research (Sun et al., 2019), it might nonetheless be worthwhile to explicitly examine the perceived helpfulness of OCRs from Black Friday purchases. To be specific, future research could, for instance, provide crowd workers a sample of OCRs that originate from both, Black Friday purchases and regular purchases, and ask them about the perceived helpfulness of the respective OCRs. With this approach, future research could further deepen the understanding of how large-scale shopping events affect OCR helpfulness in particular.

Acknowledgements

We thank the Austrian Science Fund (FWF): P32874-G for funding this research.

References

Aggarwal, P., Jun, S. Y., & Huh, J. H. (2011). Scarcity Messages. Journal of Advertising, 40(3), 19–30.

Amazon.com (2021). Amazon Announces More Deals Than Ever Before for Its 48-Hour Black Friday Event. https://press.aboutamazon.com/news-releases/news-release-details/amazon-announces-more-deals-ever-its-48-hour-black-friday-event

Bicen, P., & Madhavaram, S. (2013). Research on Smart Shopper Feelings: An Extension. Journal of Marketing Theory and Practice, 21(2), 221–234.

Blei, D. M., Ng, A.Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. Journal of Machine Learning Research, 3, 993–1022.

Byers, J. W., Mitzenmacher, M., & Zervas, G. (2012). Daily Deals: Prediction, Social Diffusion, and Reputational Ramifications. In: *Proceedings of the 5th International Conference on Web Search and Data Mining*.

Cheng, X., Yan, X., Lan, Y., & Guo, J. (2014). BTM: Topic Modeling over Short Texts. IEEE Transactions on Knowledge and Data Engineering, 26(12), 2928–2941.

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

Duan, W., Gu, B., & Whinston, A. B. (2008). Do online reviews matter? – An empirical investigation of panel data. Decision Support Systems, 45(4), 1007–1016.

Essling, I. (2020). 'Big 3' Holiday Days Exceed $22 Billion in Online Spending, Grow 25 Percent Year-Over-Year. https://www.comscore.com/Insights/Blog/Big-3-Holiday-Days-Exceed-22-Billion-in-Online-Spending

Folkes, V. S. (1988). Recent Attribution Research in Consumer Behavior: A Review and New Directions. Journal of Consumer Research, 14(4), 548–565.

Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. Cognition and Emotion, 2, 115–143.

Gelbrich, K. (2011). I Have Paid Less Than You! The Emotional and Behavioral Consequences of Advantaged Price Inequality. Journal of Retailing, 87(2), 207–224.

Grewal, D., Monroe, K. B., & Krishnan, R. (1998). The Effects of Price-Comparison Advertising on Buyers' Perceptions of Acquisition Value, Transaction Value, and Behavioral Intentions. Journal of Marketing, 62(2), 46–59.

Hayes, A. F. (2017). *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* (2nd ed.). Guilford Press.

Kalchbrenner, N., Grefenstette, E., & Blunsom, P. (2014). A Convolutional Neural Network for Modelling Sentences. In: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers*).

Lazarus, R. S. (1991). *Emotion and Adaptation.* Oxford University Press.

Li, X. (2016). Could Deal Promotion Improve Merchants' Online Reputations? The Moderating Role of Prior Reviews. Journal of Management Information Systems, 33(1), 171–201.

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

McDougall, G. H., & Levesque, T. (2000). Customer satisfaction with services: putting perceived value into the equation. Journal of Services Marketing, 14(5), 392–410.

Minaee, S., Kalchbrenner, N., Cambria, E., Nikzad, N., Chenaghlu, M., & Gao, J. (2021). Deep Learning-based Text Classification: A Comprehensive Review. ACM Computing Surveys, 54(3), 1–40.

Neuendorf, K. A. (2002). *The Content Analysis Guidebook.* Thousand Oaks.

Poniatowski, M., Neumann, J., Goerzen, T., & Kundisch, D. (2019). Organizing Their Thoughts: How Online Review Templates Affect the Review Text. In: *Proceedings of the 27th European Conference on Information Systems (ECIS)*.

Rose, R. L. (1988). Frequent-Flier Plans Become Obsessions. Wall Street Journal, September 6, B1.

Roseman, I. J. (1991). Appraisal determinants of discrete emotions. Cognition and Emotion, 5(3), 161–200.

Sánchez-Fernández, R., & Iniesta-Bonillo, M. Á. (2007). The concept of perceived value: a systematic review of the research. Marketing Theory, 7(4), 427–451.

Schindler, R. M. (1998). Consequences of Perceiving Oneself as Responsible for Obtaining a Discount: Evidence for Smart-Shopper Feelings. Journal of Consumer Psychology, 7(4), 371–392.

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. Sociological Methodology, 13, 290.

Sun, X., Han, M., & Feng, J. (2019). Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products. Decision Support Systems, 124, 113099.

Weiner, B. (1985). An Attributional Theory of Achievement Motivation and Emotion. Psychological Review, 92(4), 548–573.

Wu, J., Zhao, H., & Chen, H. (2021). Coupons or Free Shipping? Effects of Price Promotion Strategies on Online Review Ratings. Information Systems Research, 32(2), 633–652.

Zeithaml, V. A. (1988). Consumer Perceptions of Price, Quality, and Value: A Means-End Model and Synthesis of Evidence. Journal of Marketing, 52(3), 2–22.

Zhu, D. H., Ye, Z. Q., & Chang, Y. P. (2017). Understanding the Textual Content of Online Customer Reviews in B2C Websites: A Cross-Cultural Comparison between the U.S. and China. Computers in Human Behavior, 76, 483–493.

Zhu, D. H., Zhang, Z. J., Chang, Y. P., & Liang, S. (2019). Good discounts earn good reviews in return? Effects of price promotion on online restaurant reviews. International Journal of Hospitality Management, 77, 178–186.