THE POWER OF ONLINE CUSTOMER REVIEWS IN FASHION E-COMMERCE – AN EMPIRICAL ANALYSIS ACROSS CATEGORIES AND BRANDS

Jan Kemper
RWTH Aachen University, Aachen, kemper@win.rwth-aachen.de

Follow this and additional works at: http://aisel.aisnet.org/ecis2017_rp
THE POWER OF ONLINE CUSTOMER REVIEWS IN FASHION E-COMMERCE – AN EMPIRICAL ANALYSIS ACROSS CATEGORIES AND BRANDS

Research paper

Kemper, Jan, RWTH Aachen University, Aachen, kemper@win.rwth-aachen.de

Abstract

Online Customer Reviews (OCRs) have become a powerful marketing tool for e-commerce companies and an important information source for customers in online shopping. We study the influence of OCRs on sales and conversion rates of experience goods along different levels of product involvement and brand equity. We make use of a unique dataset with about 2.8 billion product detail page (PDP) views, 85.3 million sold items along 40 product categories and .9 million OCRs from a leading European online fashion company. We find that the rating valence (i.e., the average rating), has a positive influence on sales and conversion in general. Further, we identify the positive influence of OCRs on sales and conversion to be stronger for products with a high level of product involvement. Finally, this study shows that brand equity has a negative influence on the relationship between rating valence, sales, and conversion. From a managerial perspective, this study helps to use OCR as a marketing tool in the most efficient and effective way as companies should implement category and brand-specific OCR strategies.

Keywords: eWOM, Online Customer Reviews, e-Commerce, Product Involvement, Brand Equity

1 Introduction

Word of mouth (WOM) is one of the most powerful tools in marketing communication (Day, 1971; King et al., 2014). Electronic WOM (eWOM) is defined as its digital counterpart with Online Customer Reviews (OCRs) as the most common form (Jiménez and Mendoza, 2013). OCRs help consumers to make informed decisions when purchasing products online (Cui et al., 2012). Indeed, OCRs on retailer websites are for 52% of all consumers the most relevant source of information prior to purchase (Floyd et al., 2014). Their main characteristics are an unprecedented volume and reach compared to offline (Dellarocas, 2003), a viral dispersion effect across different platforms (Godes and Mayzlin, 2004) and an on-demand availability (Dellarocas and Narayan, 2007). Considering the classification in search, credence and experience goods (Nelson, 1970, 1974), OCRs matter the most for the latter category due to consumers' limited ability to verify all relevant product characteristics prior to purchase (Cui et al., 2012; Huang et al., 2009).

The e-commerce sector has experienced significant sales growth in the last decade (U.S. Department of Commerce, 2016). In 2015 online purchases contributed about 7% (12% in 2019) of total retail spending (Linder, 2015). Many companies in this prosperous field, such as Amazon or eBay, have used OCRs for more than two decades (Bloomberg Businessweek, 2009). The high level of trust and credibility which consumers usually attribute to OCRs (Bickart and Schindler, 2001) and the strong influence of OCRs on consumer purchasing decisions (Filieri, 2014) are between the main motivations for retailers to use OCRs for marketing reasons. In the literature, the impact of OCRs on sales is an area of particular research interest (Gu et al., 2012). Several studies have tried to reveal the concrete impact of OCRs on retailers' sales performance (Chevalier and Mayzlin, 2006; Cui et al., 2012; Godes and Mayzlin, 2004). However, they provide mixed findings regarding this relationship (Floyd et al., 2014; Kostyra et al., 2016). This holds especially true when we focus on the more relevant group of studies in the context of experience goods.
The majority of existing studies is either based on data from experiments with a limited sample size or on webcrawled outside-in data and heuristics based on publicly available sales rankings (Chevalier and Mayzlin, 2006; Sun, 2012). Therefore, the analysis of a large set of field data would be a significant contribution to this research area (Chen et al., 2011; Cui et al., 2012; Zhang et al., 2013). Moreover, to the best of our knowledge, there is no study that discusses the influence of OCRs on customers’ conversion behavior or the conversion rate as the percentage of visitors who have finally purchased the observed item. This is even more surprising regarding the extraordinary importance of conversion rate as key performance indicator in the e-commerce sector (Cotter, 2002). Further, little is known about the influence from OCRs on sales and conversion when controlling for different levels of product involvement (Baum and Spann, 2014; Gu et al., 2012) and different levels of brand equity (Ho-dac et al., 2013; Kostyra et al., 2016; Lovett et al., 2013).

Modern information technology enables practitioners and researchers in marketing to track consumers’ shopping behavior and satisfaction in many parts of their online shopping experience (Hinz et al., 2011). However, firms struggle to use these vast amounts of information (McAfee and Brynjolfsson, 2012) to create systematically actionable marketing insights (e.g., for improved customer service or product quality). OCRs have emerged as an important information source to evaluate products prior to purchase, which seems reasonable considering the overwhelming variety of products on many retailer websites (Kostyra et al., 2016; Li et al., 2011). However, because of the mixed findings in the existing literature, we need a better understanding of the precise impact of OCRs on customers' purchase behavior (Cui et al., 2012). This general insight is beneficial but of limited use in cases involving large assortments in online shops. Yang et al. (2012), Zhu and Zhang (2010), as well as Vermeulen and Seegers (2009), showed that retailers' OCR strategies should be specific for different product categories, brands, and customer groups. Especially product involvement and brand equity have to be considered in this context as consumers' perceived risk varies with the level of product involvement due to different search patterns (Zaichkowsky, 1985) as well as with the level of brand equity because stronger brands provide more credible signals than their weaker counterparts (Erdem and Swait, 1998). Hence, to ensure that retailers can use OCRs as a marketing tool in the most efficient and effective way, further knowledge regarding the moderating role of product involvement and brand equity, is needed (Cui et al., 2012; Jabr and Zheng, 2014; Li et al., 2013). This leads us to three research questions we aim to answer in this study: (i) Do positive OCRs increase sales and conversion of experience goods? (ii) How far does the level of product involvement affect this relationship? (iii) To which extent does brand equity influence this relationship?

We address these research gaps by developing and empirically testing a model based on actual transactions and review data. This allows us to analyze the influence of OCRs, namely rating valence what equals the average product rating, on sales as well as conversion. In addition, we add two moderator variables to the general model. First, we use the level of product involvement per product category as moderator to reveal its influence on the relationship between OCRs, sales, and conversion. Second, we use brand equity on the product level to discuss its moderating influence.

This study expands existing research in the areas of uncertainty reduction theory (Berger and Calabrese, 1975), theory of product classification (Nelson, 1970, 1974), involvement theory (Mathwick and Rigdon, 2004; Zaichkowsky, 1985) and brand signaling theory (Erdem and Swait, 1998; Montgomery and Wernerfelt, 1992). Apparel products are categorized as experience goods (Gehrt and Yan, 2004; Hong and Pavlou, 2014). Therefore, we answer the research questions by analyzing a unique data set from a European online fashion e-commerce company. Based on existing studies (Amblee and Bui, 2011; Clemons et al., 2006; Gu et al., 2012) we use different 2SLS log-log regression models for our estimation approach. To the best of our knowledge, we are the first to link a large amount of click stream data (2.8 billion product detail page (PDP) views for about 63 thousand products) and real sales data (85.3 million sold items) from online fashion with a significant amount of OCRs (> .9 million). In addition, this is the first paper discussing OCRs in the context of fashion – an industry that is regarded as one of the main growth drivers in e-commerce (BBC News, 2014).
The rest of this paper is organized as follows: In the next chapter, we present the relevant literature and develop our hypotheses based on previous results. Afterward, we discuss our data sample and the estimation approach. Then we present the results from our estimation process. Finally, we conclude by discussing the results, limitations, and areas for further research.

2 Related Literature and Hypotheses Development

2.1 Related Literature

Choosing the right product or service, online as well as offline, can be stressful. Consumers want to identify the intrinsic quality of a good based on all the available information, then purchase the product or service with the lowest transaction cost or lowest uncertainty (Hu et al., 2008; Williamson, 1979). Hence, based on early contributions from uncertainty reduction theory, consumers will engage in uncertainty reduction efforts when they lack knowledge of and experience with a product (Berger and Calabrese, 1975). Especially in the e-commerce sector, customers face a high degree of uncertainty during their purchase decisions. They cannot fully experience all aspects of a specific product or service until the final delivery. Thus, the provided product information is of high value during the purchase decision process in an online channel as customers seem to rely on peripheral quality signals more strongly than offline (Biswas and Biswas, 2004; Wood, 2001). To reduce product uncertainty, customers seek additional information like OCRs along the entire purchase decision-making process (Baum and Spann, 2014; Hu et al., 2008). User-generated recommendations like OCRs are more trustworthy and viewed as more credible than marketer-generated recommendations (Bickart and Schindler, 2001; Dellarocas, 2003). Therefore, OCRs have emerged as one of the most important information sources for customers in the e-commerce sector (Cui et al., 2012; Kostyra et al., 2016).

Products and services can be divided into three different categories: search, experience, and credence goods (Darby and Karni, 1973; Nelson, 1970). Full information prior to purchase can be acquired for search goods (e.g., digital cameras). Whereas, experience goods are characterized by attributes that cannot be known until purchase and product usage (e.g., hotel stay). Finally, credence goods cannot even be verified confidently after purchase and use (e.g., health foods). Thus, fashion products are defined as experience goods, especially in the context of e-commerce (Gehrt and Yan, 2004; Hong and Pavlou, 2014; Li et al., 2013). This kind of product type affects consumers' research behavior, use of information sources and final choice (Cui et al., 2012). In contrast to search products, experience products are evaluated more based on affective evaluative cues than on pure product attributes. However, customers are not able to assess the product experience in the online environment (Zhou and Duan, 2016). Consumers' perceived risk while purchasing online is usually greater in the case of experience goods as the appropriate evaluations are more subjective and less fact-based in comparison to search goods (Suwelack et al., 2011). Thus, credibility concerns matter more for experience goods and retailers have to provide strong signals to reduce the level information asymmetry and product uncertainty (Suwelack et al., 2011). User-generated information as OCRs can help to expand customers' information set. Therefore, OCRs play an even more important role for experience goods like fashion than for search goods (Cui et al., 2012).

Past studies show that OCRs affect consumer choices and product sales in e-commerce. However, the findings are mixed as some studies found a positive relationship between OCRs and sales (Chevalier and Mayzlin, 2006; Clemons et al., 2006; Cui et al., 2012; Dellarocas et al., 2007) while other authors failed to find a significant effect (Amblee and Bui, 2011; Duan et al., 2008). These empirical studies took place in different industries and product types like books (Godes and Mayzlin, 2004; Li and Hitt, 2008), beverages (Clemons et al., 2006), movies (Dellarocas et al., 2010; Liu, 2006), electronic devices (Archak et al., 2011) and consumer electronics (Cui et al., 2012; Ho-dac et al., 2013). The majority focused on the analysis of experience (e.g. books) instead of search goods (e.g. digital cameras). We also find mixed results in these subcategories: Amblee and Bui (2011) revealed a negative association between rating valence and sales for e-books while Chevalier and Mayzlin (2006)
were able to find a positive association in the case of books on Amazon – both experience goods. In prior research, different types of OCRs sources were discussed. So, authors used reviews mediated on third-party websites (Clemons et al., 2006), while others employed professional reviews (Basuroy et al., 2003) or OCRs from retailers’ websites similar to our approach (Sun, 2012).

Product involvement refers to the importance of a specific product from a consumer’s perspective in the purchase decision process (Richins and Bloch, 1986). Prior research has classified products as low or high-involvement based on the extent of risk perceived by consumers (Hoyer and MacInnis, 2008). Biswas and Biswas (2004) have identified performance risk, financial risk and transaction risk as the three main types of perceived risk. Performance risk covers consumers’ inability to check all product features prior to purchase. Financial risk represents the risk associated with the potential loss of money due to dishonest or criminal business partners while transaction risk is perceived considering privacy and legal issues. In this study, we focus strongly on performance risk, as our cooperation partner is an established and well-known online retailer with highly professional processes regarding logistics, payment, data security etc. The level of product involvement varies across individual consumers. However, product characteristics play a key role (Gu et al., 2012). Consumers’ information search behavior has been shown to vary for different levels of product involvement. Indeed, an increasing personal relevance regarding a product is correlated with a higher involvement (Zaichkowsky, 1985). Consequently, product involvement affects consumers’ information gathering prior to purchase and their final decision (Mathwick and Rigdon, 2004). Previous studies examining the role of product involvement in the context of OCRs are limited. Baum and Spann (2014) discussed the efficiency of OCRs in combination with recommender systems along products with low and high-level involvement condition. The authors showed that OCRs do not necessarily have to be beneficial for e-commerce players, as inconsistent recommendations do negatively influence consumers’ purchase decision. Gu, Park, and Konana (2012) looked into the impact of different external WOM sources on retailer performance. They analyzed the relative impact of external (e.g., feedback websites) and internal (retailer website) eWOM on retailers’ sales for high-involvement products. The authors found that external eWOM had a higher sales impact than internal eWOM, especially for high-involvement products. However, to date, there is no study that examines the influence of different levels of product involvement simultaneously on the relationship between OCRs and sales as well as conversion.

Kapferer (2004, p.13) defined a brand as a "shared desirable and exclusive idea embodied in products, services, places and/or experiences" and stated that "the more this idea is shared by a larger number of people, the more power the brand has". Thus, "brand equity" is a major marketing asset which can significantly help to build up a long-term buying relationship with consumers (Ambler, 2003; Christodoulides and Chernatony, 2010). However, there is no generally accepted definition of this term which covers a complex and multi-faceted concept (Christodoulides and Chernatony, 2010). Indeed, there is some consensus that brand equity denotes the added value to a product provided by its brand (Christodoulides and Chernatony, 2010; Erdem and Swait, 1998; Farquhar, 1989). Aaker (1991) identified brand awareness, brand associations, perceived quality, brand loyalty and other proprietary brand assets like trademarks as the conceptual dimensions of brand equity. Mühlbacher et al. (2016) showed that brand equity is the result of what consumers think about a brand and how they evaluate that knowledge. In addition, Feldwick (1996) identified brand equity as a measure of the strength of consumers’ attachment to the brand while Kamakura and Russell (1991) categorized brands as strong if many consumers know them and hold strong favorable associations. However, prior studies addressing OCRs and brands simultaneously are also very limited. Lovett, Peres, and Shachar (2013) analyzed the relationship between brands and WOM. The authors discussed the characteristics of a brand that stimulate WOM in the online as well as the offline setting. They found that social, emotional and functional brand elements drive WOM. Ho-dac, Carson, and Moore (2013) analyzed the role of brand equity in the context of OCRs. They revealed that brand equity is a significant moderator in the relationship between OCRs and sales. They observed that OCRs matter more for weak brands and less for strong brands. However, their empirical analysis is limited since they neither used actual sales data nor
a large sample size or wide product range. Moreover, the authors used advertising expenditures on the general brand level as a proxy for brand equity. We extend this approach by using a bottom-up brand score ranking that allows us to track brand equity on product category level for each brand.

### 2.2 Hypotheses Development

Previous literature agrees on theoretical meaning and influence of rating valence in general. Favorable opinions for a product are assumed to increase customers' purchase probability (Chevalier and Mayzlin, 2006; Clemons et al., 2006). In contrast, negative opinions are seen to discourage prospective customers (Dellarocas et al., 2007). However, the empirical findings are mixed as some studies report a significant positive effect of rating valence (Chevalier and Mayzlin, 2006; Clemons et al., 2006; Cui et al., 2012; Kopalle et al., 2009) and some do not find a significant effect (Amblee and Bui, 2011; Duan et al., 2008). Consumers can reduce their product uncertainty for search goods mainly through product information given by the retailer (Hu et al., 2008). Nevertheless, especially for experience goods, product uncertainty may remain high. OCRs are a valuable kind of user-generated content that consumers are actively seeking for to reduce their product uncertainty. Moreover, OCRs are usually presented prominently in the online shop on the top of the PDP, as in the case of our cooperation partner. Due to this web design, and based on a survey that shows that 98% of shoppers read OCRs on retailers' websites (Freedman, 2008), we assume that visitors pay attention at least to the summary statistics including the rating valence and rating volume. Hence, we argue that rating valence is a reliable indicator of overall product quality, leading to a lower level of product uncertainty and a higher level of trust. By this we assume rating valence to have a positive effect on customer choice probability in terms of sales and conversion. Therefore, our first two hypotheses are as follows:

**H1.** Rating valence is positively associated with sales

**H2.** Rating valence is positively associated with conversion rate

Marketing literature confirms that consumer information-search behaviors differ significantly with product involvement (Clarke and Belk, 1979). Product involvement refers to the perceived importance of a product from the customer perspective based on his needs, values, and interests (Richins and Bloch, 1986; Zaichkowsky, 1985). For low-involvement products, a consumer will rarely start extensive research for information or appropriate alternatives (Zaichkowsky, 1985). In contrast, consumers are willing to spend significantly more time to make the right purchase decision in cases of high-involvement products (Clarke and Belk, 1979). Thus, consumers spend a significant amount of time on research before buying such goods (Gu et al., 2012). Moreover, the level of product involvement is supposed to be higher for more expensive products and over concerns about making the best choice (Prelec and Loewenstein, 1998). Thus, making a wrong purchase decision of a high-involvement good can have larger financial implications than in the case of a low-involvement product (Gu et al., 2012).

On this basis we postulate our third hypothesis:

**H3.** Product involvement has a positive influence on the effect between rating valence and sales

Brand signaling theory tells us that uncertainty about product quality is a main risk driver (Erdem et al., 2006). Therefore, potential customers try to mitigate this risk through assessing additional signals or information prior to purchase. Previous studies underline the strong quality signal of brands (Erdem and Swait, 1998). Literature also points out that weaker brands provide less credible signals than stronger brands since stronger brands have a higher risk of losing established brand equity and potential future sales (Erdem and Swait, 1998). Thus, we assume that the signaling influence of OCRs overshadows the limited brand signal for weak brands, in contrast to strong brands that receive credible information from OCRs and the brand itself. Positive OCRs can help weak brands to substitute the lack of brand credibility that cannot be built up without significant efforts like large marketing or quality campaigns. Therefore, our fourth hypothesis is as follows:

**H4.** Brand equity has a negative influence on the effect between rating valence and sales
The conceptual framework presented in Figure 1 summarizes all four hypotheses.

![Conceptual Framework](image)

**3 Methodology**

**3.1 Data**

We were able to obtain a unique data set consisting of sales, review, and customer click stream data from a leading fashion e-commerce company selling products across the European market. The product assortment consists of clothes, shoes, and accessories for women, men, and children, including over 1,500 global and national brands as well as private labels. The sample covers the entire sales period of 2015. Similar to other studies, different parts of the original data set were excluded from further analysis to reduce complexity in the estimation procedures and to ensure statistical relevance (Heuer et al., 2015). Therefore, we include only products with more than 500 sold items per year. Based on this rule our final data set comprises 62,650 products across 40 different fashion categories. The product level in our sample set refers to all sizes and colors of one style. An example would be a white pair of sneakers by a particular brand of different sizes. In addition, the term product category refers to all products that share a comparable functionality: examples are the jeans, sneaker or shirt category. Visitors of the online shop can use the same category definition on the retailer's website to filter for specific items.

The products in our final data set split up into 746 different brands (e.g., Nike, Boss, and Converse). The total number of sold items is 85,295,524 items in the covered period. The total amount of OCRs for the discussed products equals 903,353 (with 274,483 or 30.4% from earlier periods than 2015). Clients are encouraged to give product feedback via email and have the opportunity to evaluate their items in two ways. First, via a five-star rating system known from Amazon.com and other studies (Chevalier and Mayzlin, 2006; Kostyra et al., 2016). Second, they can write free comments regarding their product experience. Both are presented on the PDP for each product in the online shop. Customers find a summary statistic on the top of the page. A more detailed overview of all ratings and comments is presented on the bottom of the PDP. Therefore, every PDP visitor is supposed to be influenced at least by the summary statistic of the rating valence and the rating volume. Furthermore, we are able to analyze the product-specific click behavior for each visitor in the online shop of our cooperation partner. Customers can use three different online sales channels (desktop, mobile website, mobile application). A PDP view is defined as one user click to access the PDP coming, for instance, from the catalog page or search engines like Google. The total amount of PDP views sums up to 2,777,422,214 in 2015.

In addition, we have access to data that provides us with information about the share of branded purchases for every brand in each product category. The brand score variable is a product-specific data item that we aggregate on a product-brand level by its average across all products from the same brand in the same product category to decrease the level of variability and to increase its force of expression.
The brand score with a value range from 0 to 1 is derived automatically by a SAP database and equals the share of items that were sold by searching or filtering for the specific brand (e.g., brand search via search engines as Google or brand filter on the retailers' webpage) prior to purchase over all sold items of this specific product in the same period. So, the brand score variable basically represents customers' brand awareness. A strong brand is defined as a brand that is known by many consumers that also hold favorable associations with this particular brand (Feldwick, 1996; Kamakura and Russell, 1991). This brand score is a promising method to measure brand image and brand equity way more precisely in comparison to existing literature as asked for specifically in prior research (Cui et al., 2012). Furthermore, our data set includes the average buying and sales price for each product in 2015. Finally, we observe the time since product launch as the number of months since the product's release date in the online shop and the date of our data collection to control for survivorship bias (Elton et al., 1996). Table 1 presents descriptive statistics for the main variables in the data set.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sales (in units)</td>
<td>1,361</td>
<td>1,838.69</td>
<td>501</td>
<td>160,085</td>
</tr>
<tr>
<td>Rating valence (1-5 stars)</td>
<td>4.27</td>
<td>.51</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Rating volume</td>
<td>14.42</td>
<td>32.98</td>
<td>1</td>
<td>3,106</td>
</tr>
<tr>
<td>PDP views*</td>
<td>1</td>
<td>1.45</td>
<td>.33</td>
<td>89.97</td>
</tr>
<tr>
<td>Conversion rate*</td>
<td>1</td>
<td>.64</td>
<td>.24</td>
<td>10.05</td>
</tr>
<tr>
<td>Sales Price (in EUR)</td>
<td>44.98</td>
<td>32.93</td>
<td>3.32</td>
<td>375.20</td>
</tr>
<tr>
<td>Buying Price*</td>
<td>1</td>
<td>.72</td>
<td>.01</td>
<td>9.33</td>
</tr>
<tr>
<td>Brand Score (value range 0-1)</td>
<td>.19</td>
<td>.11</td>
<td>.01</td>
<td>.58</td>
</tr>
<tr>
<td>Time since product launch (months)</td>
<td>15.75</td>
<td>8.33</td>
<td>5</td>
<td>88</td>
</tr>
</tbody>
</table>

Notes: For confidentiality reasons the mean values of selected variables (*) in this table are set to 1; SD, Min, and Max are proportional to these standardized values; SD = Standard deviation

Table 1. Descriptive statistics for data set (on product level)

The scope of this study is to analyze the general relationship between OCRs, sales, and conversion as well as to discuss the moderating influence of product involvement and brand equity on this general relationship. In order to analyze the moderator effect of product involvement in the estimation approach, we divide the entire sample into three different categories of product involvement (Baum and Spann, 2014; Gu et al., 2012; Mathwick and Rigdon, 2004). Simplified, the level of product involvement is a function of information-search behavior and price (Gu et al., 2012). As the level of product involvement is supposed to be higher for more expensive products (Prelec and Loewenstein, 1998), we take the average sales price per category as an indicator to split the 40 different product categories into three different levels of involvement. The final clustering from a low (Involvement 1) to a high (Involvement 3) level of product involvement for all categories is presented in Table 2.

<table>
<thead>
<tr>
<th>Involvement Cluster</th>
<th>Mean SPrice</th>
<th>SD SPrice</th>
<th>Categories</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Involvement 1</td>
<td>23.04</td>
<td>14.02</td>
<td>7</td>
<td>13,974</td>
</tr>
<tr>
<td>Involvement 2</td>
<td>44.14</td>
<td>28.86</td>
<td>21</td>
<td>35,531</td>
</tr>
<tr>
<td>Involvement 3</td>
<td>70.59</td>
<td>39.23</td>
<td>12</td>
<td>13,145</td>
</tr>
</tbody>
</table>

Notes: SPrice = Sales price; SD = Standard deviation

Table 2. Overview of Involvement Clusters

3.2 Estimation Approach

In order to analyze the relationship between OCRs and sales as well as conversion under consideration of the moderators product involvement and brand equity, we follow different studies for our estimation approach by applying a 2SLS log-log-regression model with instrumental variables (Amblee and Bui, 2011; Clemons et al., 2006; Gu et al., 2012; Jabr and Zheng, 2014). In our main model, we use
rating valence as a predictor of sales under consideration of different control variables. Therefore, Model 1 is presented by:

\[ \log(\text{Sales})_i = \beta_0 + \beta_1 \cdot \log(\text{RatingValence})_i + \beta_2 \cdot \log(\text{RatingVolume})_i \\
+ \beta_3 \cdot \log(S\text{PricePredict})_i + \beta_4 \cdot \log(\text{PDPViews})_i \\
+ \beta_5 \cdot \log(\text{TimeSPLaunch})_i + \beta_6 \cdot \log(\text{BrandScore})_i + \sum_{k=7}^{46} \beta_k \cdot \text{Category}_{ik} + \epsilon_i \]

where Sales represents the number of sold items for product i; RatingValence captures the average star rating (1-5) for product i; RatingVolume is the number of OCRs for each product i; PDPViews represents the number of PDP views for product i; TimeSPLaunch stands for the time range in month since the launch of product i in the shop and our data collection in early 2016; BrandScore captures the share of branded purchases for each product i and the dummy variable Category represents the appropriate fashion category for product i out of the 40 covered categories. Finally, we also want to control for the price of each product. However, the related literature tells us about the threat of endogeneity between sales and price due to dynamic pricing systems (Granados et al., 2012; Greene, 2003). To counteract this risk, we use buying price as an instrumental variable for sales price and run an appropriate OLS procedure to calculate the predictor SPricePredict for each product i.

In order to test the influence of OCRs on customers’ conversion behavior explicitly, we modify Model 1 by changing the dependent variable from sales to conversion rate as the ratio of sales per PDP views for each product i. In order to analyze the moderating influence of product involvement, we modify Model 1 by adding Involvement as three-stage categorical main effect and the corresponding interaction effect between Involvement and RatingValence. Finally, to analyze the role of brand equity in this context we modify Model 1 once again by adding the regular interaction effect between RatingValence and BrandScore. We omit to present any further details on these models.

### 3.3 Model Diagnostics and Mitigations

We use RStudio (version .99.891) and R (version 3.3.0) to run our calculations. The log-log regression models exhibit a reasonably good model fit and adjusted R² values between .57 up to .64. To ensure the validity of our estimations, we run several robustness checks and tests. Pairwise correlations of the main variables are presented in Table 3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. LogSales</td>
<td>1.00***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. LogRatingValence</td>
<td>.08***</td>
<td>1.00***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. LogRatingVolume</td>
<td>.55***</td>
<td>.10***</td>
<td>1.00***</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. LogSPrice</td>
<td>-.03***</td>
<td>.04***</td>
<td>.10***</td>
<td>1.00***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. LogTimeSPLaunch</td>
<td>.15***</td>
<td>.06***</td>
<td>.44***</td>
<td>.02***</td>
<td>1.00***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. LogPDPViews</td>
<td>.64***</td>
<td>.04***</td>
<td>.61***</td>
<td>.39***</td>
<td>.14***</td>
<td>1.00***</td>
<td></td>
</tr>
<tr>
<td>7. LogBrandScore</td>
<td>-.01*</td>
<td>.18***</td>
<td>-.05***</td>
<td>.41***</td>
<td>.13***</td>
<td>.05***</td>
<td>1.00***</td>
</tr>
</tbody>
</table>

* = p < .05, ** = p < .01, *** = p < .001.

Table 3. Correlation of model variables

All coefficients are substantially below the cut-off point of .8 (Kennedy, 2003). Furthermore, the variance inflation factors for all variables in our main models, as presented in Table 4, are far below the threshold of 10 (Dormann et al., 2013). Thus, we assume that multicollinearity is not a problem in our sample. In order to check for heteroscedasticity, we perform Breusch-Pagan tests for every model (Breusch and Pagan, 1979). The hypothesis of a homogeneous variance structure is rejected in all cases (p < .01). Therefore, in line with other papers (Clemons et al., 2006; Grüsschow et al., 2015), we use Huber-White robust standard errors in all four models (Greene, 2003; White, 1980). Endogeneity is a threat for empirical studies in the area of eWOM (Gu et al., 2012; Trusov et al., 2009). We counteract
this threat in four ways. First, we account for possible price endogeneity by using a cost-side instrumental variable approach for the sales price variable (Berry et al., 1995). Therein, we use a product’s buying price as an instrument (Heuer and Brettel, 2014). Second, about 30% (274,483) of all OCRs in our data sample are from 2014 or earlier. Consequently, sales transactions in 2015 have not influenced this significant amount of ratings per definition. Third, the ratio of reviewed (628,870) per sold items (85,295,524) is just about .7% in 2015. This underlines the minimal and just potential impact from sales on reviews in our data set. Fourth, we use a very large data set with about 63 thousand independent observations leading to highly stabilized estimation results and very low standard errors. Thus, we assume that endogeneity does not affect our final estimation approach severely. As an additional robustness check, we run backward selection procedures in R with no indication that we should drop variables from our selection. Finally, we check residual and QQ plots to ensure the general fit of our models (Greene, 2003).

4 Results

As a first step, we test the relationship between OCRs and sales (Model 1) as well as the light modification to examine the relationship between OCRs and conversion rates (Model 2). Table 4 reports the empirical results of the appropriate 2SLS log-log regression estimations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>β</td>
<td>VIF</td>
</tr>
<tr>
<td>Const.</td>
<td>2.74 *** (.04)</td>
<td></td>
</tr>
<tr>
<td>H1/2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LogRatingValence</td>
<td>.08 *** (.01)</td>
<td>1.07 .08 *** (.01)</td>
</tr>
<tr>
<td>LogRatingVolume</td>
<td>.14 *** (.01)</td>
<td>2.41 .14 *** (.01)</td>
</tr>
<tr>
<td>LogSPricePredict</td>
<td>-.32 *** (.01)</td>
<td>2.91 -.32 *** (.01)</td>
</tr>
<tr>
<td>LogTimeSPLaunch</td>
<td>-.16 *** (.01)</td>
<td>1.50 -.16 *** (.01)</td>
</tr>
<tr>
<td>LogPDPViews</td>
<td>.56 *** (.01)</td>
<td>2.30 -.44 *** (.01)</td>
</tr>
<tr>
<td>LogBrandScore</td>
<td>.08 *** (.01)</td>
<td>1.81 .08 *** (.01)</td>
</tr>
<tr>
<td>Observations</td>
<td>62,650</td>
<td></td>
</tr>
<tr>
<td>Adj. R²</td>
<td>.632</td>
<td>.616</td>
</tr>
</tbody>
</table>

= p < .1, * = p < .05, ** = p < .01, *** = p < .001. Notes: Huber-White robust standard errors in parentheses; Results for CG-specific variables omitted for brevity.

Table 4. Regression results Model 1 and Model 2

Model 1 reveals a highly significant and positive influence from rating valence on sales (β = .08, p < .001). Due to the chosen log-log method, all coefficients represent elasticities. The interpretation follows price elasticity from classic economic theory and is defined as the percentage change in demand because of a percentage change in price (Granados et al., 2012). All other control variables, such as rating volume (β = .14, p < .001), sales price (β = -.32, p < .001), time since product launch (β = -.16, p < .001), PDP views (β = .56, p < .001) as well as brand score (β = .08, p < .001), are also highly significant. Concluding, we find strong support for hypothesis 1 that rating valence has a positive effect on sales. We use PDP views in Model 1 as a control variable. Therefore, the results in Model 2 are almost identical to those of Model 1 – even though we change the dependent variable from Sales to ConversionRate. The reason is the used definition of conversion rate as sales per PDP views. Nevertheless, the relationship between rating valence and conversion rate is also positive and highly significant (β = .08, p < .001) which supports hypothesis 2. Like in Model 1, all used control variables are also highly significant.

As a second step, we focus on the estimation of Model 3 and Model 4 to analyze the moderating influence of product involvement as well as brand equity on the relationship between OCRs and sales. Therefore, we run two 2SLS log-log regression models. The summarized results are shown in Table 5.
We see a significant positive effect of rating valence in Model 3 (β = .07, p < .001). To test hypothesis 3, we have to focus on the appropriate interaction effects. These are modeled in R as contrast effects what enables us to run integrated significance tests to validate different levels of interaction with each other. The difference in interaction effects with rating valence between Involvement 1 and Involvement 2 (β = .05, p < .05) as well as Involvement 1 and Involvement 3 (β = .09, p < .05) are positive and significant. Thus, the influence of OCRs, namely rating valence, on sales under consideration of product involvement is significantly stronger for fashion categories with a high level of product involvement than vice versa. With this analysis, we find strong support for hypothesis 3. These results would also hold true for conversion rate as the dependent variable as we control again for PDP views in Model 3 like in Model 1. We do not show the results separately for reasons of simplicity.

In Model 4 we notice again a positive and highly significant effect of rating valence on sales (β = .02, p < .05). Moreover, we see a negative interaction effect between rating valence and brand score (β = -.05, p < .01). Thus, the influence of rating valence on sales, when considering brand equity, is significantly stronger for weak than for strong brands. This confirms hypothesis 4. As in Model 3, the results would also hold true for conversion rate as the dependent variable.

5 Discussion
This article empirically examines the general relationship between OCRs, namely rating valence, and sales as well as conversion in the context of experience goods. We further examine the moderation effect of product involvement and brand equity. Our research differs from prior studies mainly in three ways: First, by using a unique dataset from a European online fashion retailer consisting of 62,650 different products, 85.3 million sold items, 40 fashion categories, 746 brands, about 2.8 billion PDP views, and more than .9 million OCRs; second, by analyzing the influence of product involvement based on a clustering approach along several dozen product categories; and finally, we use an innovative bottom-up approach based on click stream data to approximate brand equity individually for each brand on product category level. Our analyses reveal a significant positive impact of OCRs on sales and conversion. Furthermore, we find product involvement and brand equity to be significant moderators in these relationships. In the following, we discuss important theoretical as well as managerial implications of this study and provide limitations together with avenues for future research.
5.1 Academic Contribution

This paper makes several contributions to the literature. First, by analyzing a large set of actual data we are able to discover a general positive relationship between rating valence and sales as well as conversion contributing to product uncertainty reduction theory (Berger and Calabrese, 1975). We can conclude that the higher the rating valence is, the higher the choice probability of customers of experience goods will be. Therefore, we can verify prior research that OCRs are a very powerful marketing tool (King et al., 2014). Furthermore, we contribute to existing literature by clarifying the role of rating valence in the context of OCRs as some authors have indicated a positive relationship (Chevalier and Mayzlin, 2006; Dellarocas et al., 2007; Li and Hitt, 2008) while others do not (Amblee and Bui, 2011; Duan et al., 2008). In addition, to best of our knowledge, we are the very first authors that include PDP view data as an additional variable, allowing us to discuss the influence of OCRs on customers’ conversion behavior. Moreover, this is the first study to analyze OCRs in the fashion sector, regarded as one of the main growth drivers in e-commerce (BBC News, 2014).

Second, we contribute to involvement theory and existing research in the area of eWOM by identifying a significant moderator effect of product involvement on the relationship between OCRs, sales, and conversion. Prior research has discussed different levels of product involvement like low (Clemons et al., 2006; Li and Hitt, 2008) and high product involvement (Gu et al., 2012) separately. We extend the existing literature by discussing and comparing three different levels of product involvement (low, middle, high) simultaneously in one study. Our analysis reveals that OCRs can help to drive sales and conversion for high-involvement products significantly better than for low-involvement products. This result is consistent with involvement theory indicating that consumers engage in extensive online research for expensive products with high product involvement that are associated with a higher level of product risk and vice versa (Mathwick and Rigdon, 2004; Prelec and Loewenstein, 1998). Our results highlight the importance to consider product involvement in understanding online consumer information search behavior as well as the impact on retailers’ marketing strategies.

Third, we contribute to brand signaling literature and research in the context of OCRs by identifying a significant influence of brand equity on the relationship between OCRs and sales and conversion. Our results reveal that positive OCRs have a stronger sales impact for weak than for strong brands. According to brand signaling theory, risk is driven by uncertainty about product quality (Erdem et al., 2006). As weaker brands provide less credible signals to consumers than strong brands, positive OCRs help to mitigate this deficit. This is an interesting opportunity for weak or no-name brands to overcome fears and product uncertainty by generating a significant amount of positive OCRs.

5.2 Managerial Implications

This study also makes several managerial contributions. First, we see a significant positive relationship between OCRs and sales proving the power of OCRs as an online marketing tool. Therefore, we strongly encourage responsible managers to use OCR functionalities in their online shops and to motivate customers to provide feedback via OCRs. There are several opportunities to ask customers for OCRs as reminder mailings after the product delivery date, referral pop-ups in the mobile application and while visiting the online store as well as review functionalities in the customer’s purchase history. However, retailers should be careful to motivate customers for OCRs too strongly – for example, by offering special discounts for positive ratings or even offerings to provide fake reviews. This could lead to the opposite effect, as customers might lose faith in OCRs on the retailer’s page if they no longer regard user-generated content as independent and unbiased. In fact, once OCRs become subject to managers will they run the risk to lose their strong position as a source of information which customers generally consider trustworthy and highly credible (Bickart and Schindler, 2001).

Second, category managers can also benefit from our results. We see product involvement as significant moderator in the relationship between OCRs and sales. Thus, those responsible for low involve-
ment categories like socks, scarves or gloves should not spend too much of their budget on OCRs due to the limited impact in these categories. To the contrary, managers of high involvement categories (e.g., winter coats, boots, and leather jackets) should definitely consider OCRs as powerful marketing instrument in their areas of responsibility. Therefore, they should even more encourage customers of high-involvement products to provide positive feedback and should monitor product, supplier and service quality closely to ensure a high amount of positive ratings.

Third, our results are also valuable for brand managers. Brand equity is specific for different product categories (Ho-dac et al., 2013). So, brand managers have to be aware of the individual brand equity per category. Traditional marketing campaigns are typically favored by strong brands in the appropriate categories. Regarding our results, strong brands should focus more on their existing brand equity and less on OCRs since they are not identified as significant sales and conversion driver for these brands. In contrast, weak brands could focus on a different strategy by investing, for instance, in product quality, that could generate consumer excitement in the first and positive OCRs in the second step. They can use positive OCRs to compensate their lack of brand equity. By doing so, weak brands can push their sales level as well as their level of brand equity, which is also beneficial, as seen in our results and prior studies (Ho-dac et al., 2013). This holds true especially for private labels or very young brands with limited product range and awareness. This general relationship evolves over time if the brand matures. Then a strategy shift towards a more balanced approach including OCRs and traditional marketing measures seems reasonable.

In conclusion, retailers should establish category, brand or even product-specific strategies on how to deal with OCRs in their online shop. Regarding our results, one overarching approach to using OCRs as an efficient marketing instrument cannot be seen as the most effective strategy.

5.3 Limitations and Areas for Further Research

The meaningful findings of our study notwithstanding, we naturally face some limitations. First, our empirical analysis is based on a very large data set. However, the data covers only one firm from one industry. Although many research papers use only data from one company like Amazon, the results from one company’s sales data should be validated and challenged. Replications of our study with different e-commerce players in one industry or even across different industries would be useful.

Second, we focus on the European market to contribute to the literature in the area of eWOM and OCRs in particular. However, the prosperous development in the e-commerce sector takes place in developing areas like China or India but also in developed countries like the US or major European countries as Germany (eMarketer, 2015). Consequently, as large online retailers face customers from different parts all over the world with a very heterogeneous cultural background they are eager to get deeper insights about the specific online shopping behavior. Thus, the area of cross-cultural research provides sufficient motivation to discuss the relationship between OCRs and sales along different countries, cultures or even continents (Hofstede, 2001, 1980). Such studies might be fruitful with exciting managerial and academic contributions.

Third, due to the huge amount of data, we had to use data-driven proxies to measure product involvement and brand equity. Maybe future research will be able to use well-established scales instead.

Finally, we want to underline the importance of textual parts of OCRs. Many retailers like Amazon encourage customers to give qualitative feedback to confirm their quantitative rating. On the one hand, written feedback allows consumers to gain a significantly wider range of additional information. So, they can extend their first impression from by important topics like material quality or sizing fit. On the other hand, qualitative OCRs are an excellent feedback source for retailers that need to be analyzed conscientiously as it might help retailers to improve product quality, supplier management, internal processes. Regarding the limited empirical research efforts integrating quantitative and qualitative OCRs up to now (Jensen et al., 2013), we consider this area as a fruitful path for further research.
References


Hoyer, W.D. and MacInnis, D.J. (2008), *Consumer Behavior*, 5th ed., South Western, Mason, OH.


