

## Using Online User-Generated Reviews to Predict Offline Box-Office Sales and Online DVD Store Sales in the O2O Era

Chieh Lee<sup>1</sup>, Xun Xu<sup>2</sup>, and Chia-Chun Lin<sup>3</sup>

<sup>1</sup> Yuan Ze University, Department of Industrial Engineering and Management, Taoyuan, Taiwan,  
chiehlee@saturn.yzu.edu.tw

<sup>2</sup> California State University, Stanislaus, Department of Management, Operations, and Marketing, College of Business Administration, Turlock, CA, United States, xxu@csustan.edu

<sup>3</sup> Huaying Inc., Taoyuan, Taiwan, linciach@mail.cptt.com.tw

Received 14 March 2017; received in revised form 30 December 2017; accepted 16 January 2018

### Abstract

With the rapid growth of e-commerce and social media, customers post online reviews on various online shopping websites and social media after their consumption experience, which generated the electronic word of mouth effect. In the online-to-offline era, companies are using multi-channels to increase customer demand, and the Effect generated by online users' reviews plays an important role in customer demand both online and offline. Few previous studies focused on using online user-generated reviews to forecast the demand of hyper-differentiated products online and offline, which is particularly hard to predict due to the various preferences of customers and complex relationship between factors. Via predictive global sensitivity analysis, this study uses online user-generated reviews posted on social media to predict customers' demand for hyper-differentiated products both online and offline, with an example in the film industry. We generate forecasting equations, which can successfully predict movie box office and online DVD store sales. The effectiveness and reliability of our approach are proved by the numerical studies of 22 randomly selected movies. Managers can use our method to access and analyze online users' review data and forecast future online and offline product sales.

**Keywords:** Online reviews, Demand forecasting, Global sensitivity analysis, Online and offline sales, Film industry

## 1 Introduction

Internet users and online shoppers are continuously increasing with the fast development of e-commerce. In the online-to-offline (O2O) era, companies are using multi-channels to increase customer demand and achieve better financial performance. The online marketing on social media or e-commerce platform such as Amazon have become popular than ever for both online and offline stores [22]. Internet users, and online shoppers are continuing to increase with the fast development of e-commerce. In the online-to-offline (O2O) era, companies are using multi-channels to increase customer demand and achieve better financial performance. Online sales and offline sales are more closely connected. Studies show the online activities (such as customer online reviews) generating higher online customer demand also positively influence offline sales in the O2O era [52]. Managers nowadays are seeking a technique that can utilize the abundant online customer information such as user generated reviews (UGRs) to better forecast their demands for both online and offline channels, and enhance their profitability.

Customer online reviews, which describe the customers' consumption experiences and evaluation of the products, generate electronic-word-of-mouth (E-WOM) effect. The E-WOM has become a voluntary advertising that can influence future customer demands [36]. Online shopping, unlike the brick-and-mortar store, usually involves the risks of customers having mistaken perception about the products, the risks of the actual products not matching customers' expectation, and the risks of e-retailing service quality [43], which are the results of lack of physical contacts with the products and retailers [20]. Customer online reviews effectively reduce the above risks for online shopping through introducing the details of the consumption experience, and thus enhance customers' intention to purchase [19]. In this way, customer reviews have significant value from customers' perspective, which influence future customers' purchasing behavior [21]. In addition, with the rapid development and higher popularity of O2O community, customer online reviews influence multi-channels (both of online sales and offline) sales [46], [28]. In order to find out the influence of customer online reviews on customer online and offline demand, in this study, we propose a decision analysis procedure that implements the predictive global sensitivity analysis (PGSA) to build a forecast model with online user-generated reviews as the predictor of demands for both offline box-office sales and online DVD stores.

Previous studies about online customers' reviews and customer demands have three characteristics. Firstly, most previous studies focus on the influence of online customers' reviews on customer online shopping intention or online customer demand, instead of the offline demand e.g., [37], [47]. Secondly, previous studies mostly discussed online customers' reviews on demand of homogeneous products e.g., [8], rather than hyper-differentiated products. Hyper-differentiated products are the products which can be designed and produced according to customers' requirements e.g. [11]. Due to the various preferences of customers, the desirability of hyper-differentiated products is hard to predict in terms of demand. Thirdly, most of the previous studies collect review data from online shopping / booking websites, and discuss the direct influence of these reviews on demands on those websites e.g., [7]. Few studies focus on the influence of reviews on third-party websites (e.g., social media and /or professional critic websites) on the customer demand. Thus, there is still a research gap of how to utilize user-generated reviews (UGRs) on third-party websites to predict both customer online and offline demands for hyper-differentiated products.

To fill in this research gap, through collecting data from a third-party review website, this study uses the global sensitivity analysis to explore the influence of customer reviews on the customer demand in film industry: the sales of movie box and online store DVDs, and predict the box office performance and online DVD sales from online reviews. This provides a snapshot of using the predictive global sensitivity model to predict consumers' consumption behavior of hyper-differentiated products. Hyper-differentiate products are also known as niche products, which refer to consumer products that attract customers by its unique features. These features might be drastically different from one product to another. Some examples are Movies, video games, customized tools, books, craft beers, etc.

Movies are a typical example of hyper-differentiated products. Over the past few years, the film industry has seen an accelerated revenue growth. For instance, the film industry's revenue in the U.S., which accounts for a quarter of the world's total revenue in film industry, increased from \$1.64 billion in 2002 to \$38 billion in 2015 [6]. As part of the cultural and entertainment industry, movies and movie review websites have achieved substantial growth alongside the wave of globalization since the start of new century.

In recent years, the Internet has become a major source for evaluating film quality. Via multiple discussion boards, consumers exchange product reviews, and such reviews and comments have evolved into a spontaneous advertising effect, known as E-WOM marketing of social communities. Online reviews on internet forums such as third-party review websites have become the important reference sources of customers' perceived value of a movie. These customers' perceived values generate the E-WOM effect and influence future customers' decision on watching movies in cinemas [3]. According to a *Business Next* report, in 2015, several small budget films achieved good box-office performance with low marketing budgets. One such example is *Gone Girl*, with a production cost of \$60 million, and revenues of \$167.8 million domestically, and \$201.5 million world-wide [4]. Media and professional movie reviewers highlight that the E-WOM of social communities has led to such results [13]. The online advertisement is three times more efficient than the TV advertisement. Thus, with the Internet providing easy access to information of customer perceived movie quality, the success of a movie depends not only on marketing budgets, but also on consistent and passionate narratives on social media before and after the release of the movie [27].

The main contributions of this study lie in the following aspects. First, this study establishes the link between online UGRs, and both online and offline customer demand of hyper-differentiated products in the O2O era. Second, using a predictive global sensitivity analysis, this study introduces an approach to predict future demand of hyper-differentiated products through current online UGR data. Finally, in practice, the proposed method can be easily implemented by the sales managers with minimal programming skills and a standard statistical modeling software.

The rest of this study is structured as follows: Section two reviews the relevant literature. Section three develops the model. Section four provides the numerical studies and discusses the results. Section five provides the theoretical and managerial implications. And Section six concludes the study.

## 2 Literature Review

This study discusses the influence of online customer reviews on the sales of hyper-differentiated products. In this section, we review the previous studies about customer online reviews and the generated E-WOM and previous studies about hyper-differentiated products.

### 2.1 Customer Online Reviews and E-WOM

E-commerce development is rapid [42]. A lot more customers post their online reviews after their consumption. This stimulates the study of customer online reviews and their generated E-WOM. Previous studies explore the incentives of customers posting online reviews [8]; the influence of customer online reviews on online sellers' performance [40]; and the mechanism of the influence through E-WOM, customer relationship management, and enhanced trust toward online sellers [44].

Among these, the better utilization of E-WOM and online management of customer reviews are highly focused [24]. Compared with traditional word-of-mouth effect, E-WOM has a faster spread speed, and can reflect customer perception toward the products and services in-time. In this way, E-WOM plays an important role in influencing companies' performance through increasing customer demand, corporate reputation, and advertising effects [5]. The E-WOM can be measured by the volume of online reviews [35], the extremely positive or negative reviews [52], and both of the volume and average rating ([9], [34]).

The quality of customer online reviews can be measured by variables such as helpfulness. Previous studies e.g., [8], [23] both find the evidence that the helpfulness is positively correlated with online product sales. In addition to helpfulness, our study is one of the first studies including another indicator of online review quality: the variable of number of people reading the review as one of the predictor of future demand.

Previous studies also discussed the contents and characteristics of the online reviews, which include sentiment, readability, and etc. [26]. Both customer ratings and textual comments are studied [5]. The main methodologies exploring customer online reviews and their influence used in previous studies include multiple regression e.g., [47] and text mining e.g., [33].

### 2.2 Hyper-Differentiated Products

The target of hyper-differentiated products is not only to be accepted by customers, or generate customer satisfaction, but also to arouse customer delight [17]. The sales of hyper-differentiated products are discussed by previous studies e.g., [10]. The hyper-differentiated products include movies [17] video games e.g. [52], local restaurants [51], and craft beers [10]. The sales of hyper-differentiated products are highly influenced by word-of-mouth [10].

Online channels and offline channels have demand interactions [2]. For film industry, the box sales are highly influenced by websites promotions [54] which connect the relationship between online promotion and offline sales. Online reviews of movies generate E-WOM, which increases consumer awareness, and also influences consumer evaluation of the movies and their purchase decisions [16]. Both expert critics and customers' reviews influence movie box performance to a large extent [3]. After controlling the effect of expert critics, both the volume and valence (i.e. variability) of user generated reviews significantly and positively influence the box-office performance [38].

While previous literature established the relationship between online customer reviews and online product sales, and online customer reviews and offline sales, their results have conflicts on which feature (i.e. valence, volume, and helpfulness) is most effective in demand forecasting [14]. In addition to the different features in online customer review, customers' perspectives toward the online review are also heterogeneous in different gender [1], nationalities [18], and their knowledge toward the product [53].

Our paper is different from previous literature by arguing there is no universal mechanism behind how each feature of online customer reviews influence the online and offline product sales. The influence mechanism is different even within same product category. Hence, we propose a modeling procedure that can be easily constructed and produce a reliable demand forecasting model from online customer review for both online and offline sales. Using predictive

global sensitivity analysis, our study explains the relationship between customer online reviews, and offline movie box sales and online DVD sales. In addition, we provide an approach to predict online and offline sales through our models. This helps the promotion of sales in film industry, and can be generalized to other hyper-differentiated products.

### 3 Model Development and Methodology

In this section, we first introduce the model notation and the data collection process. Then, we describe the data analysis process: structural equation modeling. To facilitate the data analysis process, we also describe the automation of the modeling process.

#### 3.1 Model Notation

Model notation can be found in Table 1.

Table 1: Notation

Item	Description
<b>Index</b>	
$i$	$i^{\text{th}}$ selected movies
$t$	the $t^{\text{th}}$ week after the movie's first release
<b>Variables</b>	
$Boxoffice_i$	weekly U.S. ticket sales of movie $i$ at week $t+1$
$V_i$	volume of movie $i$ at week $t$
$Ra_i$	cumulative rating of movie $i$ at week $t$
$Re_i$	total read per post of movie $i$ at week $t$
$H_i$	total helpfulness per post of movie $i$ at week $t$
<b>Estimate Parameters</b>	
$b_0$	intercept of structure equation
$b_i$	regression coefficients of independent variables
<b>Random Variables</b>	
$\varepsilon$	Residuals, where $\varepsilon \sim \text{Normal}(0, \sigma^2)$

Three model assumptions are as follows. First, this study has standard assumptions of structure equations: the residuals of independent variables are independently and normally distributed with mean equals zero and variance equals  $\sigma^2$ . Second, the user-generated reviews on the IMDb or Amazon are based on each user's experience. Third, the reviews are fair and are independently created. Fourth, for the online DVD sales data on Amazon, due to the data availability, based on the same assumptions from [15], [48], [49], we use the volume of customer rating as the proxy of the actual sales value.

#### 3.2 Data Collection

We randomly selected a set of recently released movies. For each movie, we collected its user-generated reviews (UGRs) from a professional critic website and an e-commerce platform, and collected box office sales from Box Office Mojo. Box Office Mojo contains rich data of weekly box office sales, and it is a website that collects, analyzes, critiques, and reports the American movies. Previous studies found that the ratings of customer reviews on the professional critic website have more significant impact on product sales [3]. We choose Internet Movie Database (i.e. IMDb) as the professional critical website to collect the professional critic and user generated review data. IMDb is the most popular and authoritative source for more than 4 million of movie and TV in the world with more than 250 million unique monthly visitors [25]. IMDb is dedicated to provide a platform for posting user-generated reviews, and is most well-known for the professional critic and user generated reviews of movies and TV shows.

IMDb allows users to post a text movie review along with a rating using ten-point Likert scale. It also allows other users to express their opinion about whether a post is helpful or not. For each selected movie, we collected its user-generated reviews from its first public release date to the last date. For each user-generated review, we recorded the username of the post, date of posting, text, rating, number of people reading this post, and helpfulness of this reviews. The helpfulness of the reviews shows number of people who think this post is helpful. IMDb measures both the number of people who read the particular review, and the number of people who vote this review helpful.

We used a visual web scraper *Webharvy*, which is a commercial software that can exact the use-generated reviews from websites. The strengths of WebHarvy lie on the fact that it can automatically scrape text and numbers from websites, and then save the scaped data into various formats. The procedures of our data collection include the following steps. First, we ran *Webharvy* software by inputting the website link into the address space of the software.

Second, we marked the data on the webpage that needs *Webharvy* automatically downloading, which include the number of read per review, the number of helpfulness of each review, customer ratings of the movie, review texts, and review date (see Figure 1). Third, *Webharvy* started download these data on each webpage (with the function of automatically crawling and extracting data from multiple pages by using the *link to the next page*). Fourth, we exacted the downloaded data, compiled them into an Excel file, and analyzed the data with Excel. In addition, as shown in Figure 1, IMDb shows for each review, X out of Y people found the following review useful. We collected the value of X (i.e., 9 in this example) as helpfulness, and the value of Y (i.e., 12 in this example) as number of people who read this post in this study.

**Mad Max: Fury Road (2015)**  
**User Reviews**  
[+ Review this title](#)

119 Reviews  
 Hide Spoilers   Filter by Rating: **7 Stars**   Sort by: **Helpfulness**

**★ 7/10** ⇒ **Customer Ratings**

**Mad Max is back**  
The Couchpotatoes **13 October 2015** ⇒ **Review Date**

What to say about a movie like this one? It's easy. If you expect a difficult dramatic story then you are not going to like it. But if you expect an action movie about a whole lot of modified vehicles trying to take over a big modified truck in the middle of nowhere, with creepy looking humans, then you are going to love it. And that's what I was going to expect. I like movies like this one where you just relax and let the action come to you. Not all the time but sometimes I do enjoy a movie like Mad Max. The amateurs of steam punk will also like this one because almost everybody in this movie is strangely looking. The filming itself was good, the script is like I said what it is. It should not have been difficult to write that script. The actors all did their job like they should. Charlize Theron and Tom Hardy are the main characters and they are good actors. Nice action movie if you are into stuff like that, otherwise just pass on this movie and watch something else.

**10 out of 18 found this helpful.** ⇒ **Helpfulness & Read of This Review**  
Was this review helpful?   | [Report this](#)

Figure 1: Online movie review on IMDb

Many online shopping websites such as e-bay and Amazon provide the option for customers to rate the product. For the e-commerce platform, we choose Amazon, since it provides similar options in IMDb for reviewers, and the volume of reviews is sufficiently large. To conduct a comparative study, we explore the relationship between user-generated review on the professional critic website (i.e., IMDb in this study) and on e-commerce platform (i.e., Amazon in this study). Note that the ratings on the Amazon are based on five-point Likert scale instead of the ten-point Likert scale on IMDb. We collected 1238 posts from IMDb and 3762 posts from Amazon.

### 3.3 Structural Equation Modeling

We implemented predictive global sensitivity analysis (PGSA) to investigate the connection between the box office sales and online DVD sales, and user-generated reviews from both Amazon and IMDb. As indicated in the previous literature e.g., [12], [32], the impact of the user-generated review on product sales varies from product to product. Therefore, it is difficult to develop a method that can capture the complexity and characteristics of the underlying the same mechanism between UGRs and product sales. In addition, there exists non-linear relationships and interaction effects between different indicators of UGRs and movie sales. Those relationships cannot be captured by linear

regression nor by exponential smoothing methods. The benefit of PGSA is to capture the key component of a complicated underlying function [45]. Therefore, to accommodate this, we aim to develop a procedure that can construct a structure equation to forecast the future box office and online DVD sales by implementing the predictive global sensitivity analysis.

Predictive global sensitivity analysis is first proposed by Wagner [45] to screen out the key factors within a complicated system. This method has been applied to areas such as inventory management, ordering and sales management, and operational hedging. Previous studies, such as [29], [30], implemented this method to construct a structural equation to represent a complicated underlying model. The method first created an extended data set of multiple non-linear combinations from the original factors. Via adding all the non-linear combination of original factors, we were able to expand the input factors to include the nonlinear factors, and the interactions items between factors. Therefore, the structure equations were able to explicitly model the nonlinear effect and the interactions of factors.

In this study, we followed the procedure in [30], but implement edit with a different application. We generated the extended data set that contains self-combinations, cross (two-way) combinations, and three-way combination for each independent variable. For self-combination, we defined the power of variable as  $\Delta$ , and thus the self-combination was defined as  $x^\Delta$ , where  $\Delta = \{-6, -5, \dots, -1, 1, \dots, 5, 6\}$ . The cross combinations, which are terms that include cross products of two different variables, are denoted by  $x^{\Delta_1}y^{\Delta_2}$ , where  $\Delta_1, \Delta_2 = \{-2, -1, 1, 2\}$ . For the three-way combinations, the powers of  $x, y$  and  $z$  are either 1 or -1; or 2 or -2 such as  $xyz^{-1}, xyz, xy^{-1}z^{-1}, xy^{-1}z, x^{-1}yz, x^{-1}y^{-1}z, x^{-1}y^{-1}z^{-1}, x^{-1}yz^{-1}$ ; and  $x^2y^2z^2, x^2y^2z^{-2}, x^2y^{-2}z^2, x^2y^{-2}z^{-2}, x^{-2}y^2z^2, x^{-2}y^2z^{-2}, x^{-2}y^{-2}z^2, x^{-2}y^{-2}z^{-2}$ .

In this way, the general model of PGSA can be reflected as follows:

$$y = b_0 + b_{1\dots}(Single\ Variable\ Items: V_i, Ra_i, Re_i, H_i) + b_{2\dots}(Non - linear\ Variable\ Items: V_i^n, Ra_i^n, Re_i^n, H_i^n) + b_{3\dots}(Interaction\ Variable\ Items: e.g., V_i^n \times Ra_i^n) \quad (1)$$

The dependent variable  $Y$  in the above regression model is the box office sales for each movie or online DVD sales for each movie in a certain week. The independent variables include three categories. The first category of independent variables includes single variable items, which include volume of movie  $i: V_i$ , cumulated rating of movie  $i: Ra_i$ , cumulated number of read per post of movie  $i: Re_i$ , and cumulated number of helpfulness per post of movie  $i: H_i$  in that week. The second category of independent variables includes non-linear variable items, which are the multi-power of single variable, such as  $Re_i^2$ . The third category of independent variables includes the interaction variable items, which are the cross-multiplication of various variables, such as  $(V_i)Ra_i^{-1}Re_i$ .

After generating the extended data items, the user can start fitting the linear regression model. The user creates the structure equation by using stepwise linear regression of the extended independent variable set against the box office and online DVD sales. While the significance levels might vary from case to case, we suggest to set the entering level of significance equals 0.05 and the leaving level of significance equals 0.10 as in previous literature e.g., [30]. In predictive global sensitivity analysis, the structure equations consider extensive non-linear relationship of the underlying function while keeping the implementation of the forecasting procedure accessible for the managers. For example, as previous literature e.g., [32] indicated, if a product receives an extremely negative review, this post affects more negatively to the product sales than the other posts. The structural equation produced by PGSA can capture this phenomenon by producing a negative term in the form of  $(Ra_i)^{-n}$ (other factors). It, thus, reflects the impact of negative reviews on product sales.

In general, as the structure equation is constructed, in the beginning of each week, the sales manager can input the last week's online customer reviews. Then, the structure equations can predict the next week's desired product sales. Furthermore, the managers may update the input data set with current week information in the end of the week and recalculate the coefficients of the structure equation. The major difference between our model and the previous studies using structure equations are that previous studies applied the predictive global sensitivity to a complex theoretical model, and used the resulting structure equation to approximate the behavior of the original model. However, in this study, we apply the PGSA to construct a structure equation for a black-box like mechanism that connects the elements of online reviews to online and offline product sales.

### 3.4 Automating the Modeling Process

In the sections 3.1, 3.2, and 3.3, we presented the details of a data analytics method of PGSA that can be implemented by Webharvy and Excel semi-automatically. To further extend the applicability of our method, we propose a pseudo code for the users who want to automate the entire modeling process via R software. Note that other commercial software such as SAS, SPSS, and Stata can also be used to implement our method.

The modeling process includes three major steps. The first step is the data collection for input factors such as independent and dependent variables. The second step is the PGSA process. And the third step is to output the desired demand forecast. The process flow chart is shown in the Figure 2. The process in Figure 2 involves the data collection step, which can be done automatically by the Webharvy after initial set up. It also includes the PGSA process, which can be done by programmable free statistical software such as R or commercial software such as SAS.

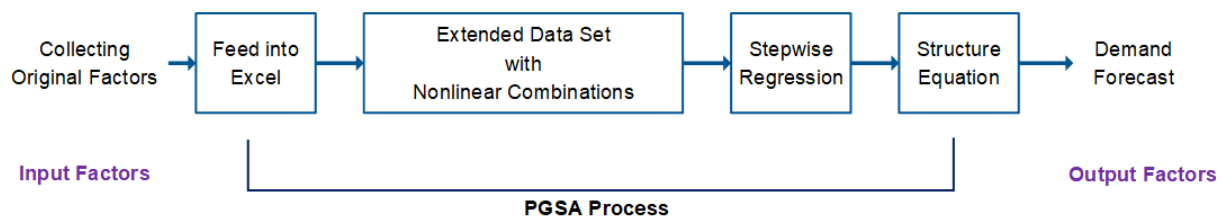


Figure 2: Process flow chart for PGSA

The code for automating modeling process is as follows:

Step 1. Setup Webharvy software to collect the input factors, and to export the collected data as an Excel file.

```

    If this is the first collection interval
    collect independent variables during the collection interval
    set independent variable = [xi]
    break, and suspend the modeling process until the next collection interval starts.
    otherwise
    collect the cumulative demand and independent variables during the collection interval
    set cumulative demand = [y]
    set independent variable = [xi]
    end;
    
```

Step 2. Input the Excel file into statistical software, and label the dependent variable as y and independent variable as xi, and merge the collected data with existing data set, if applicable. Note that for each collection interval, the new data are merged into the old data file. For example, if data collection interval is one week, and the user have implemented our method for m intervals, then y is a column vector that contains m-1 elements, and xi is a column vector that content m elements.

Step 3. Extend the input factors by adding the self-combinations and cross combinations in the data set. The combinations are calculated as follows:

```

    Input original factor = [y], [xi],
    for i = 1 to n and Δ = {-6, -5, ..., -1, 1, ..., 5, 6};
    do xadd-self = xiΔ;
    do xadd-2cross = xiΔ1 · xjΔ2, where i < j and Δ1, Δ2 = {-2, -1, 1, 2}
    do xadd-3cross = xiΔ1 · xjΔ2 · xzΔ3, where i < j < z and Δ1, Δ2, and Δ3 = {-2, -1, 1, 2}
    export data files extend data = [xi, xadd-self, xadd-2cross, xadd-3cross], and response variable = [y]
    end;
    
```

Step 4. Fit the linear regression model using statistical software.

```

    Input extend data = [xi, xadd-self, xadd-2cross, xadd-3cross], and response variable = [y]
    Set the modeling data and m = length of xi
    for all xi, xadd-self, xadd-2cross, and xadd-3cross,
    do collected the mth unit, and save them as a row vector
    set data file current week = last unit of [xi, xadd-self, xadd-2cross, xadd-3cross]
    export file current week
    end
    for all xi, xadd-self, xadd-2cross, and xadd-3cross,
    do remove the last element for each column
    set independent variable = [xi, xadd-self, xadd-2cross, xadd-3cross]
    export file independent variable
    end
    Input data files response variable = [y], independent variable = [xi, xadd-self, xadd-2cross, xadd-3cross], and current week
    = last unit of [xi, xadd-self, xadd-2cross, xadd-3cross]
    fit linear model, response = [y], and independent = [xi, xadd-self, xadd-2cross, xadd-3cross]
    Method = Stepwise, where entering level of significance α=0.05, and leaving level of significance α=0.10,
    do reserved trial current week = last unit of [xi, xadd-self, xadd-2cross, xadd-3cross]
    
```

export regression model, adj-R<sup>2</sup>, and the result of reserved trial as demand forecast end;

Step 5. Collect the output demand forecast and interpret it.

## 4 Numerical Study

We present the numerical study based on a set of randomly selected American movies released by the year 2015. We illustrate the proposed forecast model building procedure using 22 randomly selected movies. For each movie, we collect its user-generated reviews, box office sales, and online DVD sales. Because we aim to develop a procedure that can build a forecast model for a wide range of products, we went through the list of selected movie to ensure they cover the most categories of the movie. The data are collected from a 20-week interval, May 11, 2016 to September 27, 2016.

### 4.1 The Analysis of the User-Generated Reviews

We collected the user-generated review data for the 22 movies (as the names shown in Table 6). The descriptive statistics of weekly data for each movie are calculated. As an example, the descriptive statistics for variables of the Mad Max movie are shown in Table 2.

Table 2: Descriptive statistics of variables

	Mean	Medium	Max	Min	Standard Deviation
Total Number of Reviews	71.31	29.50	297.00	12.00	83.87
Customer Cumulative Ratings	493.00	188.00	2228.00	74.00	625.09
Total Number of People Read Reviews	2745.75	781.00	15575.00	273.00	3984.88
Total Helpfulness of Reviews	1312.50	325.00	7190.00	106.00	1887.27
Box Office	7172249	1246300	42651413	136899	12006423

We made the preliminary test for both data sets containing user-generated review data of 22 movies: IMDb and Amazon. We found that the volume and ratings of online reviews of professional critics from IMDb have the significantly positive correlation with the box office sales. In contrast, the volume and rating of UGRs from an e-commerce platform such as Amazon do not demonstrate the significant correlation with box office sales.

We further analyzed the review flows of these two websites, and found that the number of people reading a UGR on IMDb is much more than in Amazon. In addition, for both websites, we found that the number of helpfulness per post is highly correlated with the box office sales. Therefore, we concluded that the more appropriate approach to measure the influence per post is by adding the helpfulness into the model.

We also examined the correlation between independent and dependent variables as well as correlations between independent variables. While correlation between independent and dependent variables is high, the independent variables are also highly correlated. This is consistent with our assumption that there might be nonlinear combinations between independent variables that have significant impact on the dependent variable. Therefore, using the Predictive Global Sensitivity Analysis is essential for modeling such nonlinear relationships.

In Table 3, we present the adjusted R-square of stepwise regression model of online UGRs and box office sales for the movie of Mad Max. In the stepwise regression, we set the entering significance level equals to 0.05 and leaving significance level equals to 0.10.

Table 3: R-squares between UGRs for week *t* and Box Office Sales for week *t*+1 of Mad Max

	IMDb	Amazon
Independent Variable	R <sup>2</sup>	R <sup>2</sup>
Volume at week <i>t</i>	0.9549	0.0608
Cumulated Rating at week <i>t</i>	0.9783	0.0646
Total Read at week <i>t</i>	0.9242	0.0098
Total Helpfulness at week <i>t</i>	0.9080	0.1120



Since all four independent variables are highly correlated with the box-office sales, we use all these four variables as the input of expended data set as described in section 3.3. The sample data for the new input independent data manipulated through Excel are presented in Table 4.

Table 4: Manipulating sample new input independent variable using PGSA

Volume	Rating	Read	Helpfulness	Volume	Volume	Volume	Volume
297	2228	1085	266	88209	26198073	7.78E+09	2.31E+12
245	1797	6793	3229	60025	14706125	3.6E+09	8.83E+11
151	1027	7289	3747	22801	3442951	5.2E+08	7.85E+10
88	596	4200	2186	7744	681472	59969536	5.28E+09
49	353	1713	799	2401	117649	5764801	2.82E+08
30	199	1226	605	900	27000	810000	24300000
83	407	2905	1483	6889	571787	47458321	3.94E+09
40	268	891	372	1600	64000	2560000	1.02E+08
26	177	671	278	676	17576	456976	11881376
16	108	273	106	256	4096	65536	1048576
17	111	349	139	289	4913	83521	1419857
16	118	331	108	256	4096	65536	1048576
12	74	341	173	144	1728	20736	248832
23	149	504	226	529	12167	279841	6436343

The resulting stepwise linear regression of UGRs on IMDb against the box office sale of the Mad Max movie is

$$\text{Boxoffice}_{Mad\ Max} = -52688 - 1311.1426(V_{Mad\ Max})^2 + 239.37057(V_{Mad\ Max})(Ra_{Mad\ Max}) + 3002.0282(V)^{-2}(Re_{Mad\ Max})^2 - 11691884(V)^{-1}(Ra_{Mad\ Max})^{-1}(H_{Mad\ Max}) \quad (2)$$

For the simplicity of presentation, below we omit the subscripts of input factors in the equations, because they all share same subscript with the output variable:  $\text{Boxoffice}_i$ . In Table 5, we present the adjusted R-square of stepwise regression model of online UGRs and online DVD sales for Mad Max movie. In the stepwise regression, we set the entering significance level equals to 0.05 and leaving significance level equals to 0.10.

Table 5: R-squares between UGRs for week t and Online DVD Sales for week t+1 of Mad Max

Independent Variable	Amazon
	R <sup>2</sup>
Volume at week t	0.8518
Cumulated Rating at week t	0.8624
Cumulated Number of Read at week t	0.2224
Cumulated Number Helpfulness at week t	0.1784

The resulting stepwise linear regression of UGRs on IMDb against the online DVD sale of Mad Max movie is:

$$\text{Sales}_{Mad\ Max} = -74.6219 + 0.86485(V)(Re)(H)^{-1} + 22.23551(V)^2(Ra)^2(Re)^2(H)^2 \quad (3)$$

For hyper-differentiated products, the underlying relationship between UGRs and product sales varies from product to product [10]. Even within the same product category, the relationship can vary. Having this consideration, we manipulated the stepwise linear regression of on IMDb UGRs against box office sales and DVD sales in Amazon, with entering significance level equals to 0.05 and leaving significance level equals to 0.10. For most of the movies, the IMDb UGRs can explain over 80% of information for weekly box-office sales. For over 80% of the movies, the structure equations can explain over 85% of information of Amazon DVD sales. This implies that the e-WOM has significant impact on box-office and online DVD sales of movies. The overall adjusted R-squares for each movie are presented in Table 6. Note that for the online DVD sales of the movie of Gone Girl and Lucy, since none of the coefficients of independent variables is significant, we do not report the corresponding adjusted R-squares.

Table 6: Summary of adjust R-square values for each movie

Movie	Adjusted R <sup>2</sup> Values	
	Box Office Sales	DVD Sales on Amazon
American Sniper	0.9970	0.9697
Avatar	0.7701	0.5252
Captain America	0.9944	0.9220
Ex Machina	1.0000	0.9504
Frozen	0.7628	0.1733
Gone Girl	0.9687	N/A
Guardians of the Galaxy	0.9959	0.9385
Hotel Transylvania	0.8914	0.8205
Inception	0.9678	0.7720
Interstellar	0.5538	0.6718
John Wick	0.9917	0.6842
Life of Pi	0.9995	0.7653
Lucy	0.9039	N/A
Mad Max: Fury Road	0.8637	0.8413
Minions	0.9925	0.2262
Pacific Rim	0.9868	0.5685
Pitch Perfect	0.9895	0.7954
Sucker Punch	0.9978	0.9117
Tangled	0.9902	0.5999
The Hobbits IV	0.9986	0.3788
The Intern	0.8791	0.9144
The Secret Life of Walter Mitty	0.9988	0.9217

Using predictive global sensitivity analysis, we obtained the structure equations to predict the box-office sales by using previous week's UGRs in IMDb, which are summarized in Table 7. In addition, we obtained the structure equations to predict the online DVD sales in Amazon with previous week's UGRs in IMDb, which are summarized in Table 8. As we can observe from these equations, most of the equations contain multiple non-linear self- or cross combination of independent variables. The results are consistent with our assumption that a nonlinear relationship between independent variables and the dependent variable exists. We can also observe that structure equation for each movie is different.

Table 7: Structure equations to predict box-office sales by previous week's IMDb UGRs

Box Office Sales	Predicting Model
American Sniper	$Predicted\ Box-office\ Sale = -769213 + 15405Ra - 0.00152H^3 + (1.48E-07)(H)^4 - 23887(Ra)^{-2}(Re)^2 + 210916(V)^{-2}(Ra)^{-2}(Re)^2 + 290.0628(V)^{-1}(Re)(H) + (1.90E + 11)(V)^{-2}(Ra)^{-2}(Re)^{-2}(H)^2$
Avatar	$Predicted\ Box-office\ Sale = -13303575 + 300349(V)(Re)(H)^{-1} + (1.39E + 08)(V)^{-1}(Re)^{-1}(H) - 595.1887(V)^2(Re)^2(H)^{-2} + 5095.711(V)^2(Re)^{-2}(H)^2$
Captain America	$Predicted\ Box-office\ Sale = -814175 + 1629745(V)(Ra)^{-1}(Re) - 2944220(V)(Ra)^{-1}(H) + 372.79705(V)^2(Re)^2(H)^{-2}$
Ex Machina	$Predicted\ Box-office\ Sale = -2454843 + 126843(V)^{-1}(Re) + 10714(V)^{-2}(Ra)^2(Re)^2(H)^{-2}$
Frozen	$Predicted\ Box-office\ Sale = -7505370 + 8703.4936(V)^{-2}(Re)^2 + 46541(Ra)(Re)^{-1}(H)$

Gone Girl	$Predicted\ Box-office\ Sale = 1472451 + 428.25502(V)(Ra) - 25980(V)^2(Ra)^2(Re)^{-2} + 2858.4358(V)^2(Ra)^2(H)^{-2}$
Guardians of the Galaxy	$Predicted\ Box-office\ Sale = -1174803 - (2.1E-06)(Re)^4 + 183648(V)^{-2}(H)^2 + 4.19923(Ra)^2(Re)^2(H)^{-2} + 576723770(Ra)^2(Re)^{-2}(H)^{-2} - 9937.27(V)^2(Ra)^{-2}(H)^2 - (2.36E + 08)(V)^{-2}(Ra)^{-2}(H)^2 + 10.14301(V)^{-1}(Ra)(Re)(H)$
Hotel Transylvania	$Predicted\ Box-office\ Sale = -406103 + 480308(V)(Ra)^{-1}(H)$
Inception	$Predicted\ Box-office\ Sale = 799622 - (6.29E-09)(V)^6 - 315703(V)^{-1}(H) + 53152(V)(Ra)^{-1}(Re)$
Interstellar	$Predicted\ Box-office\ Sale = -3072493 - 7375115(Ra)^{-1}(H) + 85208(Ra)(Re)^{-1}(H) + (2.89E + 09)(V)^{-2}(Ra)^{-2}(H)^2 - 112.488(V)^2(Re)^2(H)^{-2} + 4352861(V)^2(Ra)^2(Re)^{-2}(H)^{-2}$
John Wick	$Predicted\ Box-office\ Sale = 37952 - 2.07428(H)^2 + 11529(V)^{-2}(Ra)^2 - (2.16E-05)(Ra)^2(H)^2 - 0.000852(V)^{-2}(Re)^2(H)^2 + 2.00943(V)^{-1}(Ra)(Re)(H) - 1.11014(V)^{-2}(Ra)^{-2}(Re)^2(H)^2$
Life of Pi	$Predicted\ Box-office\ Sale = 104210 + 55965(Ra) - 2867.184(V)^2(Ra)^2(H)^{-2} - 426713(V)(Re)(H)^{-1}$
Lucy	$Predicted\ Box-office\ Sale = 1349045 + (9.94E-14)(H)^6$
Mad Max: Fury Road	$Predicted\ Box-office\ Sale = -52688 - 1311.1426(V)^2 + 239.37057(V)(Ra) + 3002.0282(V)^{-2}(Re)^2 - 11691884(V)^{-1}(Ra)^{-1}(H)$
Minions	$Predicted\ Box-office\ Sale = 6910236 + 3154.6699(Ra)^2 - 12682(V)(Ra) - 9.522982(Re)(H)^{-1} + 5540.0538(V)^2(Ra)^2(Re)^{-2}$
Pacific Rim	$Predicted\ Box-office = -625777 - (2.09E-09)(Ra)^5 + 1202.9375(V)^{-1}(Ra)(Re)$
Pitch Perfect	$Predicted\ Box-office\ Sale = 36274950 - 54466378(V)^2(H)^{-2} - 419.95329(Ra)(H) - 31220174(V)^{-1}(Ra)(Re)(H)^{-1} + 43933759(V)^{-1}(Ra)(Re)^{-1}(H)^{-1} + 6902334(V)^{-2}(Ra)^2(Re)^2(H)^{-2}$
Sucker Punch	$Predicted\ Box-office\ Sale = -166139 - 105756(V) + 42750(V)(Ra)(H)^{-1} + 224619(V)(Re)(H)^{-1} + 134.65463(V)^2(Re)^{-2}(H)^2$
Tangled	$Predicted\ Box-office\ Sale = 969139 + 2410.7998(V)(Ra)(Re)^{-1}(H)$
The Hobbits IV	$Predicted\ Box-office\ Sale = 473850 - 1556198(V) - 356550(Ra) + 894006(Ra)(Re)(H)^{-1} + 615291(V)(Re)^{-1}(H)$
The Intern	$Predicted\ Box-office\ Sale = 651316 - 0.00331(V)^6 + 1431.1329(Ra)^2(Re)^2(H)^{-2}$
The Secret Life of Walter Mitty	$Predicted\ Box-office\ Sale = 138416 - 20.02625(Ra)(Re) + 0.53418(V)(Ra)(H)$

We summarize all the structural equations constructed by the PGSA for online DVD sales with the UGRs as demand predictors in Table 8. Two movies (i.e., *Gone Girl* and *Lucy*) are excluded due to the fact that none of the coefficients of predictors is significant.

Table 8: Structure equations to predict online DVD sales in Amazon by previous week's IMDb UGRs

Online DVD Sales	Predicting Model
American Sniper	$Predicted\ DVD\ Sale = 392.07855 - 1708.7402(V)^{-1} + (9.807E-05)(V)^{-2}(Ra)^{-2}(Re)^2(H)^2$
Avatar	$Predicted\ DVD\ Sale = 33.41374 - 3342.9958(Ra)^{-1} + (9.39E-25)(H)^6$
Captain America	$Predicted\ DVD\ Sale = 102.6566 - 0.01642(V)^2 + 5.28409(V)(Ra)^{-1}(H)$
Ex Machina	$Predicted\ DVD\ Sale = 190.38433 + (3.503E + 09)(Ra)^{-5} - 76.11194(Ra)(Re)^{-1}$
Frozen	$Predicted\ DVD\ Sale = 633.93251 - 1716.3732(V)^{-1}(Re)^{-1}(H) - 0.01195(V)(Ra)(Re)^{-1}(H)$
Guardians of the Galaxy	$Predicted\ DVD\ Sale = -25.55854 + 0.0577(V)^{-1}(Ra)(Re) - (6.59E-15)(V)^2(Ra)^2(H)^2 + 626.75378(V)^2(Ra)^2(Re)^{-2}(H)^{-2}$
Hotel Transylvania	$Predicted\ DVD\ Sale = 31.26654 + 1.82405(V)(Ra)^{-1}(H)$
Inception	$Predicted\ DVD\ Sale = 87.43272 - 0.00371(V)^2(Ra)^2(Re)^{-2} - 5451.109(V)^{-1}(Re)^{-1}(H) + 94667(V)^{-2}(Re)^{-2}(H)^2$
Interstellar	$Predicted\ DVD\ Sale = 64.37235 - 28792964(Ra)^{-2}(Re)^{-2}(H)^{-2}$
John Wick	$Predicted\ DVD\ Sale = 229.36248 + 2.20597(V)^2(Ra)^2(Re)^{-2}$
Life of Pi	$Predicted\ DVD\ Sale = 34.5961 + 0.16685(Ra)(Re)^{-1}(H)$

Table 8: continuation	
Mad Max: Fury Road	$Predicted\ DVD\ Sale = -74.6219 + 0.86485(V)(Re)(H)^{-1} + 22.23551(V)^2(Ra)^2(Re)^2(H)^2$
Minions	$Predicted\ DVD\ Sale = 140.62656 - 12677087(Ra)^{-6}$
Pacific Rim	$Predicted\ DVD\ Sale = 49.07076 + 2.72725(V)^{-1}(H)$
Pitch Perfect	$Predicted\ DVD\ Sale = 78.82614 - 12.05365(V)^{-2}(Ra)^{-2}(Re)^2 - 1.02245(Ra)(Re)^{-1}(H) + 17.22028(V)^{-2}(Ra)^2(Re)^2(H)^{-2}$
Sucker Punch	$Predicted\ DVD\ Sale = 4.78327 + 3.83455(Ra)(Re)^{-1} + 0.08869(V)(Re)^{-1}(H)$
Tangled	$Predicted\ DVD\ Sale = -2.83676 - 0.22333(V)^2(Ra)^2(H)^{-2} + 2.71862(V)(Re)^{-1}(H) - 0.09949(V)^2(Re)^2(H)^{-2}$
The Hobbits IV	$Predicted\ DVD\ Sale = 25.10771 + 2.19441(V)^{-1}(Re)$
The Intern	$Predicted\ DVD\ Sale = 102.79132 - 9.12655(Ra) + 176.83206(V)(Re)(H)^{-1}$
The Secret Life of Walter Mitty	$Predicted\ DVD\ Sale = 68.2239 + 0.00995(V)(Ra)(Re)(H)^{-1}$

## 4.2 Discussions of the Results

From the value of adjusted R-squares in the Table 6, we find that UGRs from the professional critic website (i.e., IMDb) can explain over 75% of variances in box office sales for 21 out of 22 movies in this study, and the value of adjusted R-squares for all movies are over 50%. In addition, from the Tables 6, we find that the UGRs from professional critic website can explain over 75% variances of online DVD sales on Amazon for 12 out of 22 movies in this study, with additional 5 movies having the value of adjusted R-squares over 50%. In this way, we find using predictive global sensitivity analysis to include non-linear items in the model can better reveal the relationship between UGRs and online and offline sales as measured by the adjusted R-squares, compared with most of the previous studies using multiple regression models e.g., [52], [39].

From Tables 7 and 8, we observe that almost all structural equations contain nonlinear terms. This result confirms our assumption that nonlinear relationships exist between UGRs, and the online and offline sales. Furthermore, our numerical study shows that other customers' opinions toward a review can serve as an indication of perceiving reliability of the online review. In other words, without sentiment analysis, the number of *helpfulness* and *read* are reliable and effective indicators for the quality and reliability of a review. Since the quality of a post can be evaluated by these two indicators, we did not incorporate the text mining into the forecasting model, and thus avoid the potential issues regarding reviews posted by non-actual customers.

Technically, from Table 7 and Table 8, we can find each of the sales forecasting model is unique, in terms of the interaction terms and their corresponding coefficients, for each movie. This reflects the strengths of predictive global sensitivity analysis, which can customize the forecasting sales model for each product to obtain a better forecasting accuracy. Although the format of each forecasting model is unique, the generation mechanism is the same through using PGSA. The terms and their corresponding coefficients can be automatically generated by stepwise regression using the code provided by Section 3.4, which facilitates the practical use by managers.

From all forecasting models in Table 7 and Table 8, we can find that most of the equations contain all of the four variables or nonlinear combination of them (i.e., *V*, volume of movie; *Ra*, cumulative rating; *Re*, total read per post of movie; *H*, total helpfulness per post of movie) examined in this study to forecast the sales. However, there are a few equations do not contain all of the four variables. For example, box office sales equation for Hotel Transylvania movie does not contain *Re* (total read per post of movie), showing the total read per post has limited role in influencing this particular movie's box sales. For another example, the online DVD sales forecasting equation for The Hobbits IV movie only contains *V* (volume of movie) and *Re* (total read per post of movie), showing the key roles of volume and total read per post in influencing the online DVD sales for this particular movie. This reflects the strengths of global sensitivity analysis on finding the key influential factors for each product sales, and also provides a guideline for managers to make the corresponding efforts to enhance the performance of these key influential factors to generate more customer demand to earn a better financial performance.

With a high success rate of constructing reliable forecasting structure equations for both offline box-office and online DVD sales, our method cannot construct an effective model for 2 out of 22 (9.09%) movies for their online DVD sales. The reason might be due to the fact that DVDs are for collection, and have more factors influencing customer purchasing decisions. Other reasons include: First, all these movies, such as Disney movies and sequel movies, already have potential buyers. Those movies more rely an established fan base. Thus, the purchasing decision is highly influenced by the previous movie instead of the current movie reviews. Second, more controversial movies, such as *Lucy* and *Gone Girl*, are widely discussed in the social media. People might watch the movie because of those positive reviews on the IMDb or the social media, but might not be interested in keeping a DVD. Third, this phenomenon might be accounted for by the fact that the potential customers tend to use the UGRs as a quality indicator of a new and/or unfamiliar product. Thus, the UGRs are much less influential to older version of movies or existing products such as the DVDs.

## 5 Implications and Future Extension

Our findings have significant theoretical and managerial implications, as discussed in this section. In addition, we describe the limitations and future research directions in this section.

### 5.1 Theoretical Implications

E-WOM has been widely studied in previous studies e.g., [31]. This study supports previous studies by confirming the significant influence of E-WOM on product sales [16]. This study focuses on the influence of online reviews on sales of hyper-differentiated products. Our finding reflects the particularly important role of E-WOM on generating the sales of hyper-differentiated products, which are customer-focused and aim to arouse customer delight. The customer demand of hyper-differentiated products we studied are in movie industry: movie box sales and online DVD sales. This bridges the E-WOM with both the online and offline customer demand in the O2O era.

E-WOM is measured by customer cumulative ratings, rating volume, total number of read, and total number of helpfulness of online reviews in weekly base in this study. This extends the previous studies' measurements about online reviews by incorporating variable of helpfulness into our models. Although previous studies found that online review plays a significant role in generating customer demand, predicting customer demand through online reviews is still vague. By utilizing predictive global sensitivity approach in this study, we successfully predict the future box sales and online DVD sales using the current periods' online review data. The comparative results obtained from third-party review websites and online shopping websites reveal the different mechanisms of online reviews' influences on hyper-differentiated products sales in online and offline channels in the O2O era.

### 5.2 Managerial Implications

Utilizing the positive E-WOM and efficient online management of reviews such as providing prompt feedback to customers are beneficial for companies' performance [24]. Managers can also utilize E-WOM to improve innovative services to enhance customer satisfaction and loyalty [41], [50] and to generate more customer demand both online and offline in the O2O era.

Since each hyper-differentiate product has its own unique characteristic [17], in this study, the predictive model of future demand for each movie is different. However, the methodology used in this study: predictive global sensitivity provides an approach for managers to generate its own demand forecasting model to predict the future customer online and offline demand for particular products.

From the perspective of implementation, the proposed forecasting model building approach is accessible for managers. Having the basic knowledge of statistics, managers can finish all the procedures and build a structure equation by using Excel with add-in functions. This provides the opportunities for managers to predict customer demand more specifically in terms of each category of products, and dynamically adjust demand forecast with incoming customer reviews.

### 5.3 Limitations and Extensions

One of the limitations of this study is that our approach might not fully reflect the critical role of negative online reviews in predicting sales [32]. Further studies can focus more on comparing the role of positive and negative reviews in forecasting product sales. Future studies can also extend the current study through the following aspects. First, more measurements from online user-generated reviews, such as score of readership and sentiment can be incorporated into the predictive model. In this way, the different influence of various dimensions of online reviews on future demand can be explored. Second, a comparative study can be conducted by finding the predictive models for both hyper-differentiated products and homogenous products. Third, the effects of moderating variables, such as actions to promote customer demand and online management of customer reviews, can be explored and taken into the predictive models. In this way, managers can better utilize the E-WOM to generate more future customer demand both online and offline.

## 6 Conclusions

As the internet becomes the main stage for people to share their opinions and to purchase their daily needs, businesses witness the enormous impact of e-word-of-mouth on both online and offline product sales. While the impact of E-WOM can be observed from various real life instances, it is difficult to be quantified. Using a predictive global sensitivity analysis approach, this study provides an accessible and effective tool for managers to utilize UGRs from professional critic websites and turn it into product online and offline demand forecasts in the O2O era.

By collecting review data from film industry, we find UGRs on the professional website significantly influence the box office sales and online DVD sales for movies. The predictive models serve as effective tools for forecasting the future demand based on the current periods of online reviews. Due to the unique characteristic of each hyper-differentiated product [17], the predictive model is different. However, through our numerical studies, we find that predictive global sensitivity analysis approach is an efficient and reliable tool to develop a demand forecasting model given by the high model fit statistics (adjusted R-square values). The predictive global sensitivity analysis can help managers to obtain the unique structural equations and provide an efficient approach for managers to forecast the future demand of hyper-differentiated products. Therefore, by implementing the proposed method, managers can effectively predict both online and offline sales with the quantitative input data from customer online reviews.

## 7 Acknowledgments

I would like to gratefully acknowledge the financial support of the Ministraton of Science and Technology of Taiwan. Technical (MOST 104-2218-E-115-005).

## References

- [1] N.F. Awad and A. Ragowsky, Establishing trust in electronic commerce through online word-of-mouth: an examination across genders, *Journal of Management Information Systems*, vol. 24, no. 4, pp. 101-121, 2008.
- [2] M. Arce-Urriza and J. Cebollada, Private labels and national brands across online and offline channels. *Management Decision*, vol. 50, no.10, pp. 1772-1789, 2012.
- [3] P. Boatwright, S.Basuroy and W. Kamakura, Reviewing the reviewers: The impact of individual film critics on box office performance, *Quantitative Marketing and Economics*, vol. 5, no. 4, pp. 401-425, 2007.
- [4] Box Office Mojo. (2016, December) Domestic Grosses. Box Office Mojo. [Online]. Available: <http://www.boxofficemojo.com/yearly/chart/?yr=2016>
- [5] A. S. Cantallops and F. Salvi, New consumer behavior: A review of research on eWOM and hotels, *International Journal of Hospitality Management*, vol. 36, pp. 41-51, 2014.
- [6] Y.Chen, S. Yang and Z. Wang, Service cooperation and marketing strategies of infomediary and online retailer with eWOM effect, *Information Technology and Management*, vol. 17, no. 2, pp. 109-118, 2016.
- [7] P. Y. Chen, S. Y. Wu and J. Yoon, The impact of online recommendations and consumer feedback on sales in *Proceedings of 25th International Conference on Information Systems*, Whashington, DC, USA,2004, p. 58.
- [8] Y .Chen and J. Xie, Online consumer review: Word-of-mouth as a new element of marketing communication mix, *Management Science*, vol. 54, no. 3, pp. 477-491, 2008.
- [9] J. A. Chevalier and D. Mayzlin, The effect of word of mouth on sales: Online book reviews, *Journal of Marketing Research*, vol. 43, no. 3, pp. 345-354, 2006.
- [10] E. K. Clemons, G. G. Gao and L. M. Hitt, When online reviews meet hyperdifferentiation: A study of the craft beer industry, *Journal of Management Information Systems*, vol. 23, no. 2, pp. 149-171, 2006.
- [11] A. S. Cui and F. Wu, Utilizing customer knowledge in innovation: Antecedents and impact of customer involvement on new product performance, *Journal of the Academy of Marketing Science*, vol. 44, no. 4, pp. 516-538, 2016.
- [12] G. Cui, H. K. Lui and X. Guo, The effect of online consumer reviews on new product sales, *International Journal of Electronic Commerce*, vol. 17, no. 1, pp. 39-58, 2012.
- [13] T. Cunningham. (2014, October) How fox marketed gone girl into a hit by keeping its secrets. *Thewrap*. [Online]. Available: <http://www.thewrap.com/how-fox-marketed-gone-girl-into-a-hit-by-keeping-its-secrets/>
- [14] P. De . Maeyer, Impact of online consumer reviews on sales and price strategies: A review and directions for future research, *Journal of Product and Brand Management*, vol. 21, no. 2, pp. 132-139, 2012.
- [15] C. Dellarocas, X. M. Zhang and N. F. Awad, Exploring the value of online product reviews in forecasting sales: The case of motion pictures, *Journal of Interactive Marketing*, vol. 21, no. 4, pp. 23-45, 2007.
- [16] W. Duan, B. Gu and A. B. Whinston, The dynamics of online word-of-mouth and product sales-An empirical investigation of the movie industry, *Journal of Retailing*, vol. 84, no. 2, pp. 233-242, 2008.
- [17] H. H. Friedman, T. Lopez-Pumarejo and L. W. Friedman, A new kind of marketing: Creating micro-niches using resonance marketing, *Journal of Internet Commerce*, vol. 6, no. 1, pp. 83-99, 2008.
- [18] J. Fong and S. Burton, Online word-of-mouth: a comparison of American and Chinese discussion boards, *Asia Pacific Journal of Marketing and Logistics*, vol. 18, no. 2, pp. 146-56, 2008.
- [19] J. Fogel and S. Zachariah, Intentions to use the yelp review website and purchase behavior after reading reviews, *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 12, no. 1, pp. 53-67, 2017.
- [20] E. Garbarino and M. Strahilevitz, Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation, *Journal of Business Research*, vol. 57, no. 7, pp. 768-775, 2004.
- [21] A. Graf and P. Maas, Customer value from a customer perspective: a comprehensive review, *Journal Für Betriebswirtschaft*, vol. 58, no. 1, pp. 1-20, 2008.
- [22] A. Gesenhues, (2015, February) Online Vs. TV: 72% of agencies say online video ads are as effective - or more effective - than TV. *Marketing Land*. [Online]. Available: <http://marketingland.com/online-vs-tv-72-agencies-say-online-video-ads-effective-effective-tv-survey-118854>
- [23] A. Ghose, P.G. Ipeirotis and B. Li, Designing ranking systems for hotels on travel search engines by mining user-generated and crowd-sourced content, *Marketing Science*, vol. 31, no. 3, pp. 493-520, 2011.

- [24] B. Gu and Q. Ye, First step in social media: Measuring the influence of online management responses on customer satisfaction, *Production and Operations Management*, vol. 23, no. 4, pp. 570-582, 2014.
- [25] E. Glassman. (2017, April) IMDb announces the top Indian movies and stars of the month as determined by page views. *Business Wire*. [Online]. Available: <http://www.businesswire.com/news/home/20170404006467/en/IMDb-Announces-Top-Indian-Movies-Stars-Month>
- [26] N. Hu, I. Bose, N. S. Koh, and L. Liu Manipulation of online reviews: An analysis of ratings, readability, and sentiments, *Decision Support Systems*, vol. 52, no. 3, pp. 674-684, 2012.
- [27] D. Kehrler. (2015, July) 8 insights on how marketing drives movie box office sales. *Forbes*. [Online]. Available: <https://www.forbes.com/sites/forbesinsights/2015/07/01/8-insights-on-how-marketing-drives-movie-box-office-sales/#4b781498ce35>
- [28] D. J .Kim, D. L. Ferrin and H. R Rao, A trust-based consumer decision-making model in electronic commerce: The role of trust, perceived risk, and their antecedents, *Decision Support Systems*, vol. 44, no. 2, pp. 544-564, 2008.
- [29] P. Kouvelis, C. L. Munson and S. Yang, Robust structural equations for designing and monitoring strategic international facility networks, *Production and Operations Management*, vol. 22, no. 3, pp. 535-554, 2013.
- [30] C. Lee and C. L. Munson, A predictive global sensitivity analysis approach to monitoring and modifying operational hedging positions, *International Journal of Integrated Supply Management*, vol. 9, no. 3, pp. 178-201, 2015.
- [31] J. Lee, J. N., Lee and H. Shin, The long tail or the short tail: The category-specific impact of eWOM on sales distribution, *Decision Support Systems*, vol. 51, no. 3, pp. 466-479, 2011.
- [32] J. Joo, Roles of the buyer's trust in seller in posted-price model of consumer to consumer e-commerce, *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 10, no. 3, pp. 30-44, 2015.
- [33] H. Li, Q. Ye and R. Law, Determinants of customer satisfaction in the hotel industry: An application of online review analysis, *Asia Pacific Journal of Tourism Research*, vol. 18, no. 7, pp. 784-802, 2013.
- [34] Y. Li, C. Wu and P. Luo, Rating online commodities by considering consumers' purchasing networks, *Management Decision*, vol. 52, no. 10, pp. 2002-2020, 2014.
- [35] Y. Liu, Word of mouth for movies: Its dynamics and impact on box office revenue. *Journal of marketing*, vol. 70, no. 3, pp. 74-89, 2006.
- [36] P. Mikalef, M. Giannakos and A .Pateli, Shopping and word-of-mouth intentions on social media, *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 8, no. 1, pp. 17-34, 2013.
- [37] J. Q. F. Neto, J. Bloemhof, and C. Corbett, Market prices of remanufactured, used and new items: Evidence from eBay, *International Journal of Production Economics*, vol. 171, pp. 371-380, 2016.
- [38] R. Niraj and J. Singh, Impact of user-generated and professional critics reviews on Bollywood movie success, *Australasian Marketing Journal*, vol. 23, no. 3, pp. 179-187, 2015.
- [39] G. Pang, F. Casalin, S. Papagiannidis, L. Muyldermans, and Y. K. Tse, Price determinants for remanufactured electronic products: A case study on eBay UK. *International Journal of Production Research*, vol. 53, no. 2, pp. 572-589, 2015.
- [40] C. Park and T. M. Lee, Information direction, website reputation and eWOM effect: A moderating role of product type, *Journal of Business research*, vol. 62, no. 1, pp. 61-67, 2009.
- [41] R. Ramanathan, An empirical analysis on the influence of risk on relationships between handling of product returns and customer loyalty in E-commerce. *International Journal of Production Economics*, vol. 130, no. 2, pp. 255-261, 2011.
- [42] I. Rodríguez-Ardura, A. Meseguer-Artola and J. Vilaseca-Requena, Factors influencing the evolution of electronic commerce: An empirical analysis in a developed market economy, *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 3, no. 2, pp. 18-29, 2008.
- [43] J. R. Rossiter, Toward a valid measure of e-retailing service quality, *Journal of Theoretical and Applied Electronic Commerce Research*, vol. 2, no. 3, pp. 36-48, 2007.
- [44] B. A. Sparks and V. Browning, The impact of online reviews on hotel booking intentions and perception of trust, *Tourism Management*, vol. 32, no. 6, pp. 1310-1323, 2011.
- [45] H. M. Wagner, Global sensitivity analysis, *Operations Research*, vol. 43, no. 6, pp. 948-969, 1995.
- [46] W. Wang, G. Li and T. C. E. Cheng, Channel selection in a supply chain with a multi-channel retailer: The role of channel operating costs, *International Journal of Production Economics*, vol. 173, pp. 54-65, 2016.
- [47] X. Xu, C. L. Munson and S. Zeng, The impact of e-service offerings on the demand of online customers, *International Journal of Production Economics*, vol. 184, pp. 231-244, 2017.
- [48] Q. Ye, R. Law and B. Gu, The impact of online user reviews on hotel room sales, *International Journal of Hospitality Management*, vol. 28, no. 1, pp. 180-182, 2009.
- [49] Q. Ye, R. Law, B. Gu, and W. Chen, The influence of user-generated content on traveler behavior: An empirical investigation on the effects of e-word-of-mouth to hotel online bookings, *Computers in Human Behavior*, vol. 27, no. 2, pp. 634-639, 2011.
- [50] J. Yu, N. Subramanian, K. Ning, and D. Edwards, Product delivery service provider selection and customer satisfaction in the era of internet of things: A Chinese e-retailers' perspective, *International Journal of Production Economics*, vol. 159, pp. 104-116, 2015.
- [51] Z. Zhang, Q. Ye, R. Law, and Y. Li, The impact of e-word-of-mouth on the online popularity of restaurants: A comparison of consumer reviews and editor reviews, *International Journal of Hospitality Management*, vol. 29, no. 4, pp. 694-700, 2010.
- [52] F. Zhu and X. Zhang, Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics, *Journal of Marketing*, vol. 74, no. 2, pp. 133-148, 2010.

- [53] P. Zou, B. Yu and Y.H. Hao, Does the valence of online consumer reviews matter for consumer decision making? The moderating role of consumer expertise, *Journal of Computers*, vol. 6,no. 3, pp. 484-488, 2011.
- [54] F. Zufryden, New film website promotion and box office performance, *Journal of Advertising Research*, vol. 40, no. 1-2, pp. 55-64, 2000.