Economic Research-Ekonomska Istraživanja

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rero20

Dynamic preference elicitation of customer behaviours in e-commerce from online reviews based on expectation confirmation theory

Keyu Lu & Huchang Liao

To cite this article: Keyu Lu & Huchang Liao (2023) Dynamic preference elicitation of customer behaviours in e-commerce from online reviews based on expectation confirmation theory, Economic Research-Ekonomska Istraživanja, 36:2, 2106275, DOI: 10.1080/1331677X.2022.2106275

To link to this article: https://doi.org/10.1080/1331677X.2022.2106275

Published online: 27 Aug 2022.

ECONOMIC RESEARCH-EKONOMSKA ISTRAŽIVANJA
2023, VOL. 36, NO. 2, 2106275
https://doi.org/10.1080/1331677X.2022.2106275

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Dynamic preference elicitation of customer behaviours in e-commerce from online reviews based on expectation confirmation theory

Keyu Lu and Huchang Liao

Business School, Sichuan University, Chengdu, China

ABSTRACT

Preference change, also known as preference drift, is one of the factors that online retailers need to consider to accurately collect consumer preferences and make personalised recommendations. Online reviews have been widely used to analyse the preference drift of consumers. However, previous studies on online reviews ignored the psychological perceptions of consumers in terms of satisfaction. This paper aims to develop a method for dynamic preference elicitation from online reviews based on exploring the theory of consumer satisfaction formation. Based on the framework of expectation confirmation theory, we develop formulas for expressing the relations among expectation, perceived performance, confirmation, and satisfaction. We then use the proposed dynamic preference elicitation model to predict the change of consumer overall preference after each review and rank products for consumers' next purchase. We test the proposed approach with a case study based on a data set from Amazon.com. It is founded that the satisfaction changes in each purchase, and this change will affect the prediction of the next product ranking. The case study is based on one product group, and further research is needed to see if the operation of the proposed method can be extended to other kinds of products.

ARTICLE HISTORY

Received 22 December 2021
Accepted 21 July 2022

KEYWORDS

Behavioural OR; preference elicitation; preference drift; online reviews; consumer satisfaction

JEL CODES

C80; D12; M30

1. Introduction

According to the report by eMarketer (2021),[1] retail ecommerce sales worldwide climbed to $4.213 trillion in 2020 and will exceed $7 trillion in 2025. The rapidly increasing online sales gives online retailers a glimpse of potential huge profits, which stimulates them to try all kinds of methods to grab the attention of customers. It is a common practice of online retailers to elicit users' preferences regarding personalised services based on historical data. Preference elicitation helps online retailers learn users' preferences and then recommend matching products to the users, thus achieving increasing sales ultimately. In addition, online customers can also benefit from

CONTACT Huchang Liao liaohuchang@163.com

preference elicitation, such as obtaining a list of favourite products. Therefore, preference elicitation is of great practical value for increasing the efficiency of online sales.

As it is noted by Bernstein et al. (2018), online retailers need to collect an abundance of customer data for personalise sales. Online reviews are one of the most important data sources reflecting consumers' preferences, and have been used in many studies to analyse consumer behaviours (Park & Lee, 2021; Shen et al., 2019). From online reviews, product attributes that consumers value can be mined, which, to some extent, can be regarded as a process of consumer preference elicitation. However, such a process ignores psychological perceptions of consumers. Existing studies on consumer preference elicitation based on online reviews rarely considered satisfaction theory, which may affect the accuracy of preference elicitation. In fact, we can not only extract consumers' preferences, but also obtain consumer satisfaction from online reviews. As one of the important factors in consumer purchase decision studies (Dash et al., 2021; Li et al., 2020), consumer satisfaction can be used to help to recommend or exclude products. Existing studies on consumer satisfaction mainly focussed on factors of satisfaction and satisfaction's influence on repurchase intention (Han & Hyun, 2018; Hult et al., 2019). In addition, influenced by age increasing, culture environment, social interaction and other factors, the preferences of users for different products change both in the short term and in the long term (Fehr & Hoff, 2011; Fung & Carstensen, 2003; Pereira et al., 2018). The timely elicitation of users' dynamic preferences contributes to precisely obtain the preference information of customers, which is helpful for improving the accuracy of product recommendation (Štefko et al., 2019).

To solve the problems that the current studies about preference elicitation ignored the psychological perceptions and preference changes of consumers, this paper aims to develop a method for dynamic preference elicitation from online reviews based on exploring the theory on consumer satisfaction formation. Expectation confirmation theory (ECT) is a classic theory about the causes of satisfaction proposed by Oliver (1980), which holds that consumer satisfaction is jointly determined by consumer expectation and the degree of confirmation between expectation and perceived performance. The ECT has been widely used in different contexts (Filieri et al., 2021; Lee & Kim, 2020), which can help to understand the post-adoption purchase behaviour of consumers. This paper utilises the ECT to explain the formation of consumer satisfaction. Note that most studies on the ECT conducted hypothesis tests on the factors of satisfaction and the influence of satisfaction on repurchase intention based on the information provided by consumers (Liao et al., 2017; Shin et al., 2017). However, collecting consumer data on site can hardly meet the requirements of the big data era. It is not practical to collect the expectation, perceived performance and satisfaction information of every online customer during each purchase in the form of questionnaires.

In this paper, based on the purchase and review behaviour of a target customer, we define his/her expectation and perceived performance with the online reviews of other consumers and himself/herself, respectively. In addition, we propose formulas to express the relations among expectation, perceived performance, confirmation and satisfaction. Then, we propose a dynamic preference elicitation model where retailers use online reviews of customers as the source of preference elicitation. After observing a customer purchase and comment on products in the same product category, we use the proposed model to elicit his/her preference reflected in the purchase. The ECT is used to calculate the consumer satisfaction, in which we define each element in ECT using formulas based on online reviews, and ultimately measure consumer satisfaction per purchase. When updating consumer preferences, the effect of different degrees of satisfaction on dynamic preference elicitation is taken into account, which involves the change of consumer satisfaction. Finally, we propose a product ranking method based on the proposed dynamic preference elicitation model. The contributions of this paper are as follows:

1. We redefine the expectation and perceived performance in the ECT based on online reviews, and propose a method to estimate customer satisfaction without direct interactions with customers. Our work can help to adapt the traditional ECT to the context of online shopping.
2. To improve the accuracy of product recommendation, we propose a dynamic preference elicitation model considering consumer satisfaction, in which every online review and changing satisfaction of consumers are used to update preferences.
3. We illustrate the practical value of the dynamic preference elicitation model in a realistic setting by using a data set from Amazon.com. From the case study, it is verified that our dynamic preference elicitation model is efficient and implementable.

The rest of this paper is organised as follows: Section 2 provides a review of the relevant literature. Section 3 introduces the proposed model, while Section 4 illustrates its effectiveness through a case study. The paper ends with conclusions in the final section.

2. Related work

This paper dedicates to proposing an approach to learn dynamic consumer preferences from online reviews based on the ECT. Consumer satisfaction has an impact on the effect of each extracted preference to the overall preference of a product in the process of dynamic preference elicitation. Thus, to facilitate further presentation, this section reviews relevant work related to consumer satisfaction, the ECT, and the changes of preferences. At the end of each section, we put forward corresponding improved methods to circumvent the deficiencies in existing literature.

2.1. Consumer satisfaction and the ECT

Consumer satisfaction is a managerial research topic with enduring popularity. Although many studies regarding satisfaction have been conducted (Yuksel & Yuksel, 2001), there is not a consensus on the definition of satisfaction (Giese & Cote, 2000). Oliver (1980) defined that consumer satisfaction is a degree to which the *expectation*

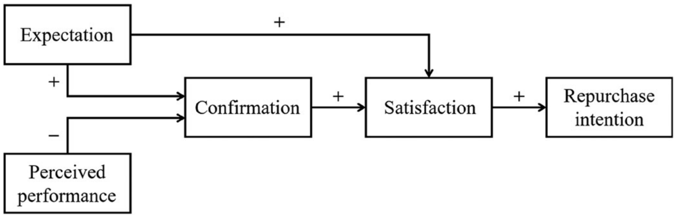

Figure 1. Relationships between factors in the expectation confirmation theory.
Source: Cited from Bhattacherjee (2001).
Note. + and − denote positive relation and negative relation, respectively.

matches the perception of a consumer, which was the most widely used definition of satisfaction (Prayag et al., 2019). Since satisfaction is an individual subjective evaluation, it was considered to be closely related to the psychology of people. However, the thinking process of a consumer was like a black box, which meant that an observer could only see satisfaction results, without knowing how it was generated (Oliver, 2014). In this regard, Helson (1959) collected data in the form of questionnaires and built structural equations to measure the relations between the antecedents and consequences of satisfaction. The analysis results showed that consumer satisfaction was a function of the *expectation* level before shopping and the extent to which the expectation was confirmed. In addition, the satisfaction of a consumer also influences his/her *repurchase intention*. To interpret the causes of consumer satisfaction in a comprehensive manner, Oliver (1980) proposed a cognitive model which came to be known as the ECT. To facilitate understanding, the relationships between factors in the ECT can be intuitively demonstrated in Figure 1, which has been illustrated in Bhattacherjee (2001).

The relation between satisfaction and repurchase intention is one of the most popular research issues in marketing. The positive impact of satisfaction on repurchase intention has been proved in Mittal and Kamakura (2001) and Liao et al. (2017). After the original ECT was proposed, different influence factors of satisfaction and repurchase intention of consumers were added into the expectation confirmation model, and their reliabilities were assessed by hypothesis test. For example, with regard to the factors of satisfaction, trust was taken into account in Shin et al. (2017). Given that the ECT is an effective consumer satisfaction analysis tool, the positive effect tendency between high satisfaction and high repurchase intention has been proved in many papers regarding the ECT (Li et al., 2020). However, to the best of our knowledge, the intensities of effect of different degrees of satisfaction on subsequent repurchase decisions has not been investigated in the ECT, which consists an issue that we will consider in this study.

We consider that high satisfaction levels of products and services reflect the liking of consumers. We can learn consumer preferences by collecting the satisfaction data of consumers regarding each purchase. In early days, scholars usually collected data in the form of questionnaires and interviews to conduct empirical studies on satisfaction. With the developments of technology, natural language processing towards online consumer reviews was applied to analyse consumer satisfaction (Godnov & Redek, 2019; Liu et al., 2021). In this study, we use the consumer information extracted from online reviews as the data source. Although consumers' perceptions on products can be extracted from online reviews, how consumer expectation is obtained from online reviews has not been studied. Therefore, how to identify *consumer expectation* information from online reviews so as to obtain *consumer satisfaction* according to the ECT is worth thinking about. After determining consumer satisfaction, we can recommend customised products and services to customers according to their *repurchase intentions*.

2.2. Preference drift

The preferences of users may change owing to short-term reasons such as the seasonality or periodicity of some products. For example, consumers' preferences regarding refrigeration equipment will increase in summer. There are also long-term preference changes. For instance, users' preferences will change as they age and their living environment changes. For instance, old people pay more attention to emotional meaning with regard to advertisements (Fung & Carstensen, 2003).

As research continues, the feature that preference changes with time, also known as preference drift, has attracted much attention and many methods have been proposed to deal with the preference drift problem. Users' preferences can be learned based on historical data. Koren (2010) pointed out that because data changed over time, models should be updated to accurately interpret the information contained in the data. Rafeh and Bahrehmand (2012) utilised a time decay function to predict the behaviours of users: the closer the rating time of users to the current time is, the greater the function value is. Instead of completely excluding old records, forgetting functions can consider old and new information simultaneously. Different forgetting functions were used in Zheng and Li (2011) and Feng et al. (2015) to express the long-term preference drift of users. Sahoo et al. (2012) proposed a hidden Markov sequence to explain the changing product selection behaviours of users over time. Shi (2014) used the timestamp information to filter the ratings of products given by other users related to the current user, and the ratings and products outsides the timestamp would not be referenced.

For products and services in the same store, regular consumers will face different situations for different purchases, and their satisfaction will also change. Although various methods, like the expectation confirmation model (Oliver, 1980) and servqual model (Parasuraman et al., 1988), have been used to measure satisfaction, the dynamics of satisfaction were ignored. Therefore, it is necessary to measure satisfaction from online reviews within a dynamic setting. Observe that most of the above-mentioned methods used proposed functions based on time to assign a weight to preferences at each moment. We consider that it is not necessary to assign weights exactly according to time, but applying certain methods to combine the emerging preference with the old one can also reflect the change of preferences.

3. Proposed approach

In this section, we first formalise the consumer preference elicitation problem. In particular, we define the notion of consumer preference. Then, we propose a dynamic preference elicitation model based on online reviews in Section 3.2, where we explain the elements of the proposed model, i.e., consumer expectation and consumer satisfaction in detail in Section 3.2.1 and Section 3.2.2, respectively. Finally, we propose a product ranking method using the proposed dynamic preference elicitation model.

3.1. Problem definition

On comprehensive online shopping websites such as Amazon (www.amazon.com), Taobao, (www.taobao.com) and eBay (www.ebay.com), there are many kinds of products which have diverse attributes. It is not rational to recommend one type of products to a consumer based on the consumer's preferences for another type of products. For example, we cannot predict that a customer prefers furniture with little formaldehyde according to his/her preference for sweet food. In order to improve the accuracy of preference learning, it is necessary to classify products, just like Aliexpress (www.aliexpress.com), all products on which are grouped into 13 categories such as Women's Fashion, Consumer Electronics, and Jewellery & Watches. Preference learning regarding products in different categories is independent. For instance, if a consumer reviews a cookie on an online platform, his/her diet preferences will be learned, but his/her preferences for other types of products, such as clothing, could not be obtained from this online review on the cookie. In this study, we suppose that there are mutually exclusive categories on an online platform containing all products on sale, which can be interpreted clearly in Figure 2, where each product category is a slice.

In this study, the preference elicitation is based on different slices. We consider a customer who browses a shopping or service website as a user. Let user x browse a website to make a purchase, and $B = \{1, \ldots, b\}$ denote a set of purchase decisions of x. We use $t \in B$ to denote the current purchase behaviour of x, while $t-1 \in B$ denotes that last purchase and $t+1 \in B$ denotes the next purchase. Consider an online platform with n_t items in a product category corresponding to the purchase behaviour t, and the set of all products is denoted as $N_t = \{1, \ldots, n_t\}$. Suppose that a user buys product $y \in N_t$ in his/her purchase t. For $y \in N_t$, we suppose that there are l_t reviewers, the set of which is described as $L_t = \{1, \ldots, l_t\}$. If one consumer reviews a product multiple times, we treat those reviews as if they were from different consumers. Since this study is based on online reviews, if a consumer does not make

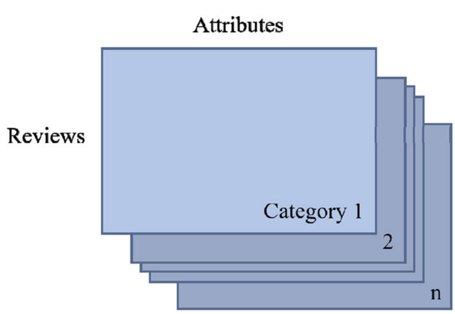

Figure 2. Preference elicitation based on different product categories.
Source: Created by the authors.

any comment after a purchase, then, this purchase behaviour will not be taken into account.

Definition 1 (Consumer preference). Suppose that there are m attributes for products in a category. The consumer preference is defined as the weights given by a consumer to product attributes. $W'^x_t = \{w'^x_{1,t}, \ldots w'^x_{m,t}\}$ denote attribute weights extracted from the $t-$th online review of x, which are regarded as the $t-$th set of sub-preferences. $W^x_t = \{w^x_{1,t}, \ldots, w^x_{m,t}\}$ denote the set of overall preferences of x composed of the former t sets of sub-preferences.

Note that the overall preference will change with each updated online review.

3.2. Dynamic preference elicitation model

In this section, we propose a dynamic preference elicitation model based on the ECT. Figure 3 illustrates the proposed model, which shows the process of eliciting preference of consumers from online reviews.

In the left side of Figure 3, we divide the input online reviews into two parts, including the reviews of other consumers and the review of the user. The own review is exclusively the $t-$th review of the user on a kind of product, while the review(s) of others are reviews made by other consumers before the user made the $t-$th purchase decision. Specific data collection, preparation and evaluation will be detailed in the following sections.

The middle of Figure 3 is the data processing part of our dynamic preference elicitation model. Through a natural language processing, online reviews are converted into the weights and sentiment scores that consumers assign to each product attribute, which will be introduced specifically in Section 3.2.1. We first calculate the expectation, perceived performance, and sub-preference of the user from online reviews. It is noted that the expectation of the use is determined based on the online

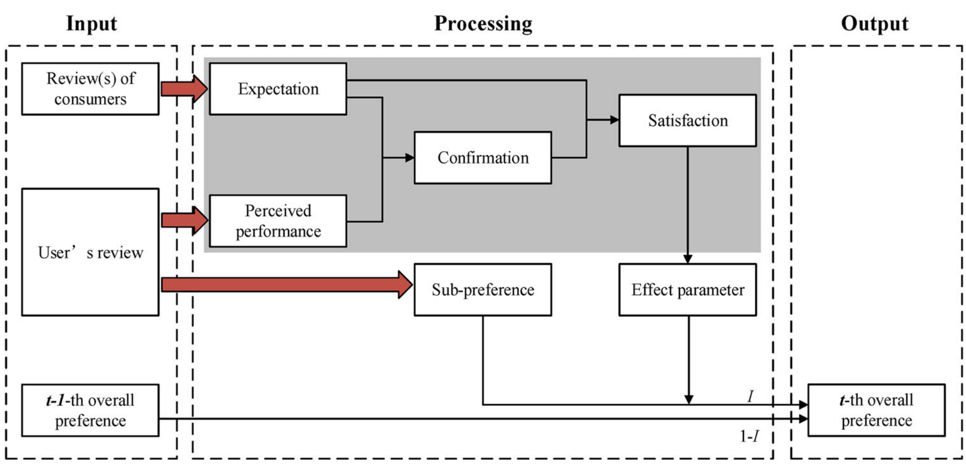

Figure 3. Framework of the dynamic preference elicitation model.
Note. The shaded area is ECT. I is the effect parameter value, reflecting the influence degree of sub-preference on the $t-$th overall preference.
Source: Created by the authors.

reviews of other consumers. According to the ECT, we compute the confirmation of expectation and perceived performance, and then calculate the satisfaction based on the expectation and confirmation. Note that the classic ECT depends on the data directly provided by consumers collected in the form of questionnaires, while in our study, the expectation, perceived performance, and satisfaction are all redefined based on online reviews. Motivated by the conventional ECT that the satisfaction will influence repurchase intention, in this study, we suppose that the satisfaction will affect the influence degree of each sub-preference to the overall preference. Then, combining the sub-preference in the $t-$th round and the $(t-1)-$th overall preference, we can get an updated overall preference in the $t-$th round, in which the user satisfaction is taken into account as a mediating factor. In what follows, we describe each part of our dynamic preference elicitation model in details. It is noted that the consumer expectation and consumer satisfaction is indeed the expectation and satisfaction of the user in the $t-$th round of purchase.

3.2.1. Consumer expectation

The expectation of individuals is subjective. In other words, the expectation of someone is not known until it is expressed verbally. Against the backdrop of online consumption, when a user gets to a website, sellers do not have access to the user's expectation towards a variety of products and his/her preference.[2] However, an important advantage of online platforms, compared with traditional offline consumption, is that previous consumers' reviews are publicly displayed on platforms. It is helpful for those who want to quickly learn the function, quality and other features of a product to refer to the online reviews of others. Many studies (Mauri & Minazzi, 2013; Sparks et al., 2016; Zhang et al., 2014) have proved that online reviews have a significant impact on the expectation and purchase intention of consumers. In this subsection, user expectation will be represented by the information extracted from past consumers' online reviews.

Different consumers usually concentrate on different attributes of the same product. For example, people with high incomes may be insensitive to price and thus the price, as a product attribute, will not appear in their reviews; but it may be an important attribute for low income consumers. In addition, different consumers may evaluate the attributes of the same product differently. For instance, what is good for the skin of some people may cause skin allergies in others. In general, owing to personalised preferences, different reviewers usually have different comments regarding the same product. For each reviewer, his/her preferences can be regarded as the weights of product attributes, and the attribute weights of different reviewers are diverse. Sentiment analysis, also known as opinion mining (Li et al., 2021), detects opinions and attitudes of users according to the elements of textual information. Through sentiment analysis methods such as machine learning and lexicon-based approaches, a large number of user-generated reviews can be processed (Montejo-Ráez et al., 2014; Soumya & Pramod, 2020; Zhao et al., 2018). We use the tool of sentiment analysis to extract the weights of different product attributes given by different consumers, which consists of following three steps:

***Step* 1 (Text preprocessing)**. We first use regular expression algorithms[3] to clean the raw data, including removing dates and urls in reviews that are not valuable for analysis. Stop words, which are extremely common and carry little text information, like demonstratives and conjunctions, are also removed.

***Step* 2 (Attribute extraction).** Latent dirichlet allocation (LDA) (Blei et al., 2003), as a text analysis method which is appropriate for unstructured online reviews, is applied to cluster words and extract the main product attributes. Let $M = \{1, \dots, m\}$ denote the set of m extracted attributes of products.

***Step* 3 (Computing attribute weights).** Let the weights of attributes in the t−th purchase be $W_t^j = \{w_{1,t}^j, \dots, w_{m,t}^j\}$, $j \in L_t$. In the results of Step 2, each product attribute contains a number of sub-attributes. For example, in hotel services, park, car, and street all belong to the attribute of parking. For reviews that do not mention any sub-attribute, each attribute is equally weighted. For reviews that mention sub-attributes, if all of the sub-attributes of an attribute are not mentioned, the weight of this attribute is zero. The weight of each attribute can be determined according to the proportion of words in its category. For the review given by reviewer $j \in L_t$, the proportion of attribute h can be calculated by

$$pro_{h,t}^j = \frac{f_{h,t}^j}{f_{\text{all},t}^j}, \ \ h \in M \tag{2}$$

where $f_{h,t}^j$ is the number of sub-attributes belonging to the attribute h, and $f_{\text{all},t}^j$ denotes the number of all words in the review of j. The weight $w_{h,t}^j$ is the outcome of normalisation of $pro_{h,t}^j$. In our study, if a word appears many times in a review, we count each time.

After determining the weights of attributes, we then need to aggregate the preferences of all past consumers from online reviews to get consumer expectation. Each online review has an impact on different users. In general, a user tends to focus on the reviews that mentioned the attributes they are interested in. For example, price-conscious consumers are more sensitive to words like 'cheap' and 'worthless'. Therefore, to aggregate all reviews as a component of consumer expectation, different reviews need to assign different weights, and reviews with attribute weights close to those of a user are given a large weight.

We use the Euclidean distance to measure the difference between the preferences of each reviewer and a user. Let $W_{t-1}^x = \{w_{1,t-1}^x, \dots, w_{m,t-1}^x\}$ denote the attribute weights of user x, that is, his/her original overall preference obtained according to his/her previous purchases. If the customer is new, his/her preference is expressed as equal weight for all attributes. The distance $d_t(j,x)$ between the attribute weights of reviewer j and user x is calculated by

$$d_t(j,x) = \sqrt{\sum_{h=1}^{m} (w_{h,t}^j - w_{h,t-1}^x)^2}, \ \ j \in L_t \tag{3}$$

Then, the weight of reviewer j can be calculated by

$$v_{j,t}^{x} = \frac{\frac{1}{d_t(j,x)}}{\sum_{k=1}^{l}\frac{1}{d_t(k,x)}},\ j \in L_t \tag{4}$$

where $0<v_{j,t}^{x}<1$ for $j \in L_t$ and $\sum_{j\in L} v_{j,t}^{x} = 1$.

We define adjectives, adverbs and verbs, for example, *happy*, *strenuously*, and *pain*, that usually express the sentiment of a sentence, as sentiment terms. By assigning sentiment scores to those words, for example, '*+2*' is given to '*very good*' and '0' is assigned to '*medium*', we can convert each review into attribute ratings corresponding to each consumer. We pre-process online reviews of consumers, where all sentences are divided into words which are then clustered. Let $A_t^j = \{a_{1,t}^j, \ldots, a_{m,t}^j\}$ denote the sentiment scores of m attributes extracted from the review of reviewer j, $j \in L_t$ and let $a_{h,t}^j \in [1,5]$, that is, the positive sentiment score at the extremely positive level is 5 and the most extremely negative sentiment score is 1. For example, in an online review '*this is the worst service I have ever encountered*', the superlative of the adjective is used, and accordingly, we give the service attribute a score of 1. Stanford CoreNLP,[4] a natural language processing tool with many functions such as word segmentation and stemming, is used in this study to obtain the sentiment score of each attribute.

For a user in online shopping, online reviews are the main source of consumer expectations towards a product when the specifics of the product are unknown. The acceptance of online reviews is influenced by the preference of the user. In other words, a user pays more attention to their preferred product attributes and reviews that are more similar to his/her overall preference. We can define the aggregated preferences of past consumers as the expectation of a user as follows:

Definition 2 (User expectation). For user x in the purchase t, his/her expectation is obtained by

$$E_t^x = \sum_{j=1}^{l}\sum_{h=1}^{m} v_{j,t}^{x} w_{h,t-1}^{x} a_{h,t}^{j},\ t \in B \tag{5}$$

where $v_{j,t}^{x}$ is the weight of reviewer $j \in L_t$, $w_{h,t-1}^{x}$ is the $(t-1)-$th sub-preference of consumer x, and $a_{h,t}^{j}$ is the sentiment score of attribute $h \in M$ given by reviewer $j \in L_t$. According to the range of $a_{h,t}^{j}$, it is easy to obtain that $E_t^x \in [1,5]$.

3.2.2. Consumer satisfaction

Some websites like www.jd.com, allow consumers to give an overall star rating for a product, which to some extent can be interpreted as *consumer satisfaction*. However, it is possible that different consumers have different sentiments regarding the same star rating. For instance, customers with high tolerance may still give high overall star ratings in the face of low-quality products, while grumpy consumers may give low start ratings due to slow logistics without considering the quality of products. There will be deviation from the actual evaluation of consumers, if the overall star rating is used to determine the satisfaction of consumers, which will further lead to a decline in the accuracy of consumer preference prediction and product recommendation. In the setting of this study, the expectations and perceived performances of consumers in the ECT are all derived from online reviews, which cover more information and can represent the actual sentiments of consumers more accurately. Therefore, the ECT based on online reviews is utilised to calculate the consumer satisfaction in this section.

Before the specific calculation method of consumer satisfaction is proposed, we first define the perceived performance and confirmation in the ECT. Consumers will form a real perception of a product after they bought and used it. This perception is directly reflected in the online reviews of consumers. According to the reviews written by x for the purchase t, we can extract the corresponding attribute weights, also named the $t-$th sub-preference, which are denoted by $W'^x_t = \{w'^x_{1,t}, \ldots w'^x_{m,t}\}$, and the $t-$th sentiment scores are denoted by $A^x_t = \{a^x_{1,t}, \ldots, a^x_{m,t}\}$. We define the perceived performance as

$$P^x_t = \sum_{h=1}^{m} w'^x_{h,t} a^x_{h,t}, \quad t \in B \tag{6}$$

Similar to E^x_t, the range of P^x_t is also $[1, 5]$.

Then, we define the confirmation of expectation and perceived performance of x in the purchase t as

$$C^x_t = \frac{P^x_t}{E^x_t}, \quad t \in B \tag{7}$$

where $C^x_t \in [0.2, 5]$.

According to the ECT, the consumer expectation, and the confirmation of expectation and perceived performance, have positive influence on consumer satisfaction. In this paper, consumer satisfaction is regarded as a function of the expectation and confirmation, which is denoted as:

$$S^x_t = \alpha^x_t E^x_t + \beta^x_t C^x_t \tag{8}$$

where $0 < \alpha^x_t, \beta^x_t < 1$, $\alpha^x_t + \beta^x_t = 1$. Based on the range of E^x_t and C^x_t, there is $S^x_t \in [0.2, 5]$. For the sake of narrative convenience, in the following, α and β are used to denote the coefficients of the expectation and confirmation, which are also defined as satisfaction parameters.

It is worth noting that the effect of satisfaction on subsequent purchases is not proportional. For example, it is obvious that, compared with products with satisfaction of 0.2, customers are more likely to repurchase products with satisfaction of 5. As for the product with satisfaction of 2.6, the consumer is not dissatisfied with it, but the probability that he/she will buy the product in the next purchase is vague. In other words, a mid-level satisfaction rating for a product hardly reflects consumer preferences. In previous studies, the asymmetric effect of online reviews has been discussed. For example, Forman et al. (2008) pointed out that, compared with extreme

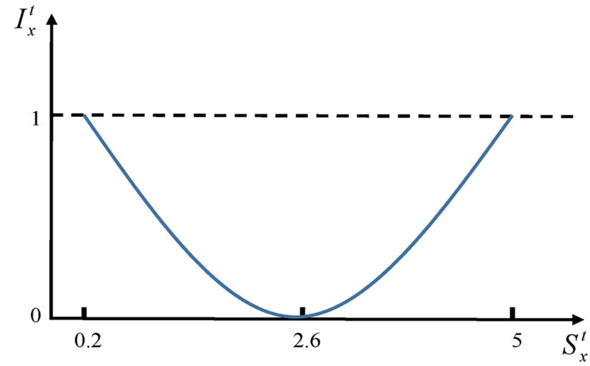

Figure 4. Sample of the relation between satisfaction and effect.
Source: Created by the authors.

ratings, moderate ratings were considered less helpful. Pavlou and Dimoka (2006) also proved that the extremely positive or negative evaluation on online sellers provided more information than moderate ratings. In summary, extreme evaluation information, such as very high or very low satisfaction, is more useful than medium satisfaction for future prediction.

Let I_x^t denote the effect of the $t-$th sub-preference on the overall preference of x, which changes according to the value of each calculated satisfaction, as shown in Figure 4. Let $I_t^x \in [0, 1]$, where '0' means that the overall preference will not change in this purchase under the satisfaction condition, and '1' means that the sub-preference information extracted in this purchase will be fully utilised when calculating the overall preference. For example, as can be seen in Figure 4, when $S_t^x = 2.6$, there is $I_t^x = 0$, which can be interpreted as that the preference change of x cannot be judged due to the little information generated in the online review of purchase t.

Previous studies like Forman et al. (2008), only proved the strong usefulness of extreme ratings, while the specific function has not been proposed to fit the effect under different degrees of evaluation. Taking Figure 4 as an example, we have determined the coordinates (0.2,1), (2.6,0) and (5,1), but the shape of the line between these three coordinates is unknown. Park and Nicolau (2015) tested the U-shape relation between star rating and its effect, but the accurate shape was still unknown, which can only be obtained by analysing historical data of consumers. Through the calculation of the real data on online platforms, not all satisfactions of consumers can cover the whole interval [0.2,5], that is, it is impractical to collect the complete satisfaction data of every consumer to learn the corresponding relation between the satisfaction and effect. In addition, the impact of satisfaction on effect is also variable. If we match the values of satisfaction and effect accurately, there may be over-fitting problems, which will reduce the accuracy of preference elicitation. Thus, this paper only roughly takes into consideration the convex relation between satisfaction and effect. By using the *smoothing spline curve fitting* tool of MATLAB, we plot the effect of the sub-preference extracted in each purchase on the overall preference, under the effect of satisfaction, as shown in Figure 5.

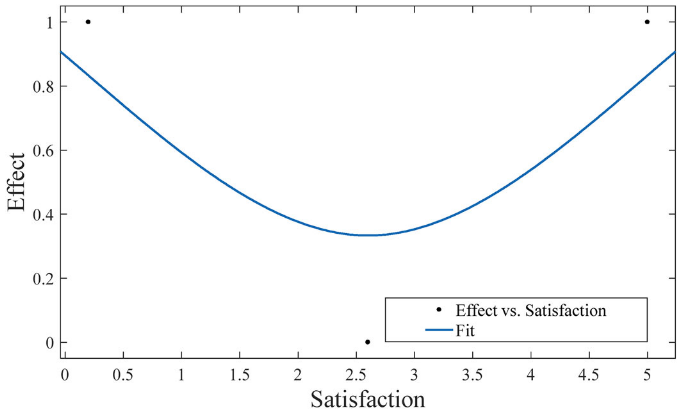

Figure 5. The fitting relation between satisfaction and effect.
Source: Created by the authors.

We regard the effect of sub-preference on overall preference as the weight of sub-preference. In this sense, the $t-$th overall preference can be computed by

$$W_t^x = I_t^x \cdot W_t'^x + (1 - I_t^x) \cdot W_{t-1}^x \tag{9}$$

To sum up, after a customer reviews a product online, consumer satisfaction can be calculated by Equation (8). Then, according to the fitting curve in Figure 5, the corresponding value of effect can be obtained, in other words, the weight of sub-preference extracted from each online review can be calculated. Finally, the overall preference can be updated according to (9).

3.2.3. Determining satisfaction parameters

Product ranking is based on consumer expectation for each product. In order to derive the expectations of products for ranking, we need to discuss the dynamics of satisfaction parameters α and β given in Equation (8).

The degrees to which satisfaction is influenced by the expectation and confirmation vary from person to person and are changed over time, which leads to the need to pay attention to the changes of α and β. In Section 3.2.2, we set satisfaction as a necessary factor to compute overall preference. However, the parameters α and β change with each review. We cannot calculate the real overall preference without knowing the values of these two parameters, and thus cannot make personalised product recommendations for consumers. Therefore, we use predicted values of these parameters to make recommendations, and then use the user's following purchase decision as a feedback to determine the real value of these parameters, the details of which is explained below.

From Oliver (1980), we can learn that the impact of confirmation on satisfaction is much greater than that of expectation on satisfaction. Therefore, in the calculation of α and β which are satisfaction parameters of expectation and confirmation respectively, to make sure that α is greater than β and $\alpha + \beta = 1$, we suppose that α varies in [0,0.5] and the corresponding range of β is [0.5,1]. In this paper, each calculation of satisfaction uses predicted values of α and β based on historical data. Generally speaking, the distribution of the impact of the expectation and confirmation on satisfaction is stable, without seasonal and periodic changes. Therefore, moving average is adopted in this study to predict α and β of each time. By the moving average method, we can get the predicted parameters $\hat{\alpha}$ and $\hat{\beta}$, which will be used to calculate the predicted satisfaction $\hat{S}_t^x$, so as to get the predicted overall preference $\hat{W}_t^x = \{\hat{w}_{1,t}^x, \ldots \hat{w}_{m,t}^x\}$ according to Equations (8) and (9).

Let l_z, $a_{h,t,z}^{j_z}$ denote the number of online reviews on product $z \in N_t = \{1, \ldots, n_t\}$, and the sentiment score of reviewer j_z for attribute h of product z, respectively. According to Equation (5), the predicted expectation for product z is calculated as:

$$\hat{E}_{t,z}^x = \sum_{j_z=1}^{l_z} \sum_{h=1}^{m} v_{j_z,t} \hat{w}_{h,t}^x a_{h,t,z}^{j_z}, \quad t \in B, \quad z \in N_t \tag{10}$$

The higher the predicted expectation to the product z, the higher its rank is. Given that the numbers of products and reviews are constantly changing, we calculate the predicted expectation of a product on a daily basis and change product rankings accordingly.

For new consumers, we use $\alpha = \beta = 0.5$ as default values. The real values of α and β are obtained as follows:

***Step* 1 (Enumeration)**. List all combinations of α and β with a step size of 0.1 such that $\alpha \in [0, 0.5]$, $\beta \in [0.5, 1]$, $a + \beta = 1$. It is obvious that there are 6 groups of α and β. Let $G = \{1, \ldots, 6\}$ denote six difference kinds of combinations of α and β.

***Step* 2 (Predicting overall preference).** Based on the expectation and confirmation for each purchase, we use all combinations of α and β to calculate satisfaction separately. Then, calculate the effect of sub-preference extracted from the online review of a purchase on the overall preference. Finally, predict the overall preference.

***Step* 3 (Optimal coefficients judgement).** To find the optimal coefficients, we rank all products according to their different expectations $E_{t,z}^x z \in N_t$ calculated by Equation (10) based on different overall preference. Note that Equation (10) contains the overall preference of each consumer, so the corresponding product ranking is different for different consumers, where personalised service for customers is reflected.[5] In Step 2, we obtain different overall preferences using different α and β. Therefore, for consumer x, there are 6 kinds of ranking. Compared with other rankings, if the selected product in the purchase t has the highest position in the ranking obtained by the combination $g^* \in G$, we regard α and β in g^* as the optimal parameters. The values of these two parameters will be used to calculate the overall preference and next predicted parameters.

3.3. Product recommendation

To facilitate implementation, in this section, we summarise the process of personalised product recommendation for users based on preference elicitation. The steps of personalised recommendations based on our dynamic preference elicitation method are as follows:

***Step* 1.** Extract expectation, perceived performance and sub-preference from online reviews, and compute confirmation based on expectation and perceived performance by Equations (5) and (6).
***Step* 2.** Compute predicted satisfaction using the moving average method based on expectation and confirmation by Equation (8).
***Step* 3.** Compute predicted overall preference based on sub-preference, last overall preference, predicted satisfaction by Equation (9).
***Step* 4.** Compute the predicted expectation to each product based on predicted overall preference by Equation (10), and rank products according to predicted expectations.
***Step* 5.** Compute real satisfaction and overall preference according to the next reviewed product.

4. Empirical study

In this section, we discuss the results of empirical experiments conducted on a data set from Amazon.com. Figure 6 shows an example of the online information on Amazon.com. Considering our research is based on the products of one category and products purchased frequently are suitable for the study of dynamic preference elicitation, we selected a cosmetic as our study subject. We crawled the review data under the foundation category in BB Facial Creams department on October 12, 2021. The data set consists of 15,615 records for a set of 124 products, each of which contains product ID, reviewer ID, review text, and review time. We preprocessed those data, including removing the data without review text and review time, and retaining the data whose review text is English. Finally, there are 14,339 records belonging to 86 products remained as the basis of this study.

We select a user J who has made 4 purchases of foundation on Amazon.com as an example for analysis and one piece of his/her relevant data is shown in Table 1. Let his/her initial overall preference be $W_0^J = [1/m, \ldots 1/m]$, where m is the number of product attributes.

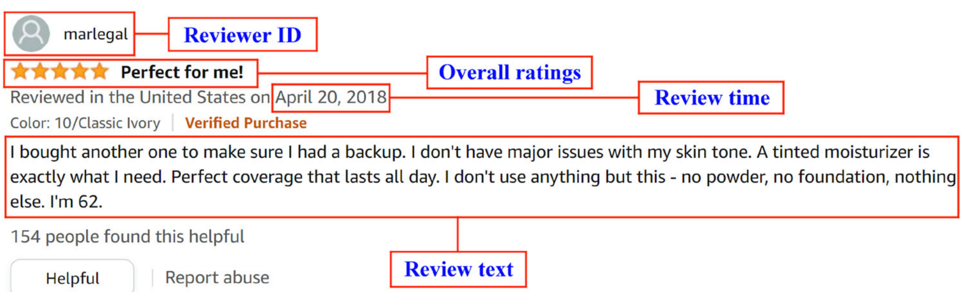

Figure 6. An example of online review on Amazon.com.
Source: Cited from Amazon.com and edited by authors.

Table 1. An example of the data of user J.

Reviewer ID	Product ID	Review time	Review text
J	B00AIEJF68	2018/9/10	This summer I was on the hunt for an affordable bb cream with spf that evened out my skin. Well this was not it. Yes it had spf 30 but past that it is just awful! I've tried all different ways of applying it, but unfortunately it just sits on my skin. It doesn't absorb well at all. It gets cakey and gives off a look of rubber (for lack of a better word) sitting on your skin. This is not a first impression I've used this many times before I decided to review it and declutter it. Don't waste your money they're better ones out there

Source: Cited from Amazon.com and edited by authors.

Table 2. Attribute extraction results.

Attribute	Keywords
Brand	brand(s), clinique, elf, L'Oréal, mac, maybelline, milani, missha, neutrogena, nyx
Color	ashy, beige, blush, brown, color, colour, dark, golden, gray, green, light, nude, orange, pale, pink, red, tan, tone(s), skintone, undertone(s), white, yellow
Function	acne, acne-prone, aging, blemish(es), bright, control, freckles, glow(ing), greasy, moisturizer(s), moisturizing, oil(y), pimple(s), pore(s), protection, rosacea, scars, sheer, spf, summer, sunscreen, tinted, winter
Packaging	bottle(s), package, packaging, porcelain, pump(s), tube(s)
Price	affordable, cheap(er), cost, expense, expensive, fair, inexpensive, money, price, value, worth
Quality	allergic, consistency, ingredients, mineral(s), quality, texture
Smell	fragrance, scent, smell(s)
Other	advertised, brush, delivery, sample, seller, service, shipped, shipping, size, sponge

Source: Created by the authors.

4.1. First preference elicitation of user J

In this section, we present the detailed preference elicitation process of user J based on his/her first review.

Before calculating the expectation, we need to extract product attributes. First, we preprocess the reviews. We use regular expression algorithms to clean the raw review texts, mainly including removing newline, date, website address, special symbols and individual letters. In addition, we remove stop words that appear frequently in online reviews but have nothing to do with cosmetics. In order to improve the accuracy of attribute extraction, we conduct frequency statistics on all words in reviews, and add the high frequency words that do not belong to any attribute into the stop words list. Then, we use the LDA (Latent Dirichlet Allocation) to extract attributes. Note that because the results of LDA will contain noisy words and repetitive topic words, the final analysis results need to be manually adjusted (Zhang et al., 2021). After merging semantically similar words and removing noisy words, we get the main attributes of foundation and their related words which can provide a basis for sentiment analysis, as shown in Table 2.

From Table 2, we can find that consumers mainly focussed on eight attributes of foundation: *Brand*, *Color*, *Function*, *Packaging*, *Price*, *Quality*, *Smell*, *Other*. Then, we can let the initial overall preference of user J be $W_0^J = [1/8, \ldots 1/8]$. We first rank the products whose review time is earlier than that of user J according to W_0^J, which is a recommendation for the target customer. It is worth noting that although online retailers recommend products for consumers according to their learned preferences, it is quite normal for consumers to select products without high ranks. After all, human preferences are complex and changeable to be completely and accurately described.

For the first product that user J reviewed, we study its 203 reviews published by other consumers before user J's purchase. We calculate the weight of each attribute in each review, the distance between the preference reflected in each review and that of user J, and the weight of each review according to Equations (2)–(4). Then, we use Stanford CoreNLP to compute sentiment scores. Assign 1, 2, 3, 4, 5 to '*very negative*', '*negative*', '*neutral*', '*positive*', '*very positive*', respectively. The results of Stanford CoreNLP are the sentiment analysis implemented based on every sentence. Therefore, in order to get the sentiment score of each attribute of each review, we average the score of each sentence. Note that not all reviews mention every attribute. For reviews that do not mention any attribute, we assign their overall sentiment scores computed by Natural Language Toolkit (NLTK), a natural language processing library of Python, to each attribute. For reviews that do not completely cover all attributes, we assign '*neutral*' value 3 to attributes that are not mentioned. Finally, based on Equation (5), we get the expectation of consumer J before his/her first purchase, that is $E_1^J = 2.924$.

According to the first review of user J, we can obtain his/her sentiment score on each attribute. Then, the first perceived performance P_1^J and confirmation C_1^J can be calculated by Equations (6) and (7), respectively. The results are $P_1^J = 2.834$, $C_1^J = 0.971$.

In Section 3, we mention that the parameters α and β change over time. We cannot get the real parameters until consumers' second product review. Therefore, product ranking for consumers after their first review is based on predicted α and β. For new consumers, let $\hat{\alpha}_1 = \hat{\beta}_1 = 0.5$. There are prediction results that $\hat{S}_1^J = 1.948$ and $\hat{W}_1^J = [0.077, 0.269, 0.221, 0.125, 0.077, 0.077, 0.077, 0.077]$.

Next, we calculate the real parameters of the first review. Based on the time of the second review, we extract all previous reviews which belong to 37 products. Under different values of α_1^J and β_1^J, on the basis of E_1^J and C_1^J being known, we obtain 6 satisfaction values and corresponding overall preferences. Then, we calculate the expectations of user J for each product under different overall preferences according to Equation (10), and rank the products accordingly. As shown in Table 3, the product reviewed by consumer J for the second time, which is labelled 32, ranks highest when $\alpha_1^J = 0$ and $\alpha_1^J = 0.1$. Considering that there is a higher expectation on the product when $\alpha_1^J = 0.1$, we determine that $\alpha_1^J = 0.1$ and $\beta_1^J = 0.9$ are real parameters for S_1^J. Correspondingly, the satisfaction and overall preference extracted from the first review are calculated as $S_1^J = 1.166$ and $W_1^J = [0.027, 0.33, 0.262, 0.125, 0.057, 0.057, 0.057, 0.057]$, respectively. In order to rank products in the next stage, we use the moving average method to predict the $\hat{\alpha}_2^J$ and $\hat{\beta}_2^J$, the results of which are $\hat{\alpha}_2^J = 0.3$ and $\hat{\beta}_2^J = 0.7$.

4.2. Collective preference elicitation results of user J

User J has made 4 reviews in the product category of foundation. In this section, we elicit preference according to each review behaviour. The preference elicitation

Table 3. Product ranks based on expectations.

	$\alpha_1^J = 0$		$\alpha_1^J = 0.1$		$\alpha_1^J = 0.2$		$\alpha_1^J = 0.3$		$\alpha_1^J = 0.4$		$\alpha_1^J = 0.5$	
Rank	$E^J_{1,PN}$	PN	$E^J_{1,PN}$	PN	$E^J_{1,PN}$	PN	$E^J_{1,PN}$	PN	$E^J_{1,PN}$	PN	$E^J_{1,PN}$	PN
1	4.000	29	4.000	29	4.000	29	4.000	29	4.000	29	4.000	29
2	4.000	30	4.000	30	4.000	30	4.000	30	4.000	30	4.000	30
3	3.988	18	3.988	18	3.987	18	3.987	18	3.987	18	3.987	18
4	3.865	28	3.871	28	3.878	28	3.883	28	3.888	28	3.893	28
5	3.729	25	3.740	25	3.751	25	3.761	25	3.771	25	3.780	10
6	3.712	10	3.728	10	3.743	10	3.757	10	3.769	10	3.779	25
7	3.620	12	3.623	12	3.629	12	3.637	12	3.646	12	3.655	12
8	3.552	11	3.559	11	3.565	11	3.570	11	3.573	11	3.575	11
9	3.496	23	3.502	17	3.507	17	3.511	17	3.513	17	3.514	17
10	3.495	17	3.501	23	3.503	23	3.504	23	3.504	23	3.502	23
11	3.477	27	3.486	27	3.493	27	3.498	27	3.500	27	3.502	27
12	3.436	19	3.443	19	3.449	19	3.453	19	3.455	19	3.455	19
13	3.419	15	3.425	15	3.429	15	3.434	13	3.439	13	3.454	1
14	**3.407**	**32**	**3.417**	**32**	3.427	13	3.430	15	3.436	1	3.443	21
15	3.404	13	3.417	13	**3.424**	**32**	**3.429**	**32**	**3.431**	**32**	3.442	13
16	3.403	37	3.411	37	3.417	37	3.421	37	3.430	15	**3.432**	**32**
17	3.391	35	3.398	35	3.402	35	3.414	1	3.427	21	3.429	15
18	3.364	16	3.365	9	3.388	1	3.408	21	3.425	37	3.427	37
19	3.351	22	3.365	16	3.386	21	3.402	35	3.403	5	3.416	5
20	3.349	9	3.364	22	3.379	9	3.391	9	3.401	35	3.408	9
21	3.338	21	3.363	21	3.374	22	3.386	5	3.400	9	3.398	35
22	3.333	8	3.357	1	3.365	5	3.382	22	3.388	22	3.391	22
23	3.324	1	3.341	5	3.364	16	3.360	16	3.360	20	3.370	20
24	3.315	5	3.333	8	3.333	8	3.346	20	3.355	16	3.350	16
25	3.301	20	3.313	20	3.329	20	3.333	8	3.333	8	3.333	8
26	3.251	26	3.255	26	3.258	26	3.280	6	3.300	6	3.315	6
27	3.200	36	3.230	6	3.257	6	3.273	36	3.292	36	3.308	36
28	3.198	6	3.226	36	3.251	36	3.259	26	3.260	26	3.260	26
29	3.182	4	3.192	4	3.201	4	3.207	4	3.223	7	3.238	7
30	3.166	24	3.152	7	3.180	7	3.203	7	3.212	4	3.225	2
31	3.152	34	3.149	24	3.136	2	3.171	2	3.200	2	3.225	3
32	3.121	7	3.136	34	3.136	3	3.171	3	3.200	3	3.215	4
33	3.058	2	3.099	2	3.129	24	3.107	24	3.085	24	3.064	24
34	3.058	3	3.099	3	3.118	34	3.099	34	3.080	34	3.062	34
35	2.742	33	2.731	33	2.720	33	2.711	33	2.702	33	2.694	33
36	2.000	14	2.000	14	2.000	14	2.000	14	2.000	14	2.000	14
37	2.000	31	2.000	31	2.000	31	2.000	31	2.000	31	2.000	31

Note. PN = product number. The product reviewed by user J and corresponding expectation are highlighted in bold.
Source: Created by the authors.

Table 4. Expectation, perceived performance, confirmation and sub-preference of each review.

t	E^J_t	P^J_t	C^J_t	W'^J_t
1	2.924	2.834	0.971	[0,0.5,0.375,0.125,0,0,0,0]
2	3.119	2.583	0.828	[0,0.833,0,0.167,0,0,0,0]
3	3.031	2.7	0.891	[0,0,0.6,0,0.4,0,0,0]
4	3.044	2.576	0.846	[0,0.727,0.182,0,0.091,0,0,0]

Source: Created by the authors.

process on the first review is shown in Section 4.1 and we omit the detailed steps of other reviews in this section.

In Tables 4–6, we present results computed based on 4 reviews of user J. Table 4 shows expectation, perceived performance, confirmation and sub-preference extracted from each review, in which expectation and confirmation are used to compute real and predicted satisfaction. Table 5 presents the values of α, β, satisfaction and overall

Table 5. Real α, β, satisfaction and overall preference of each review.

t	(α_t^J, β_t^J)	S_t^J	W_t^J
0	(0.5,0.5)	Null	[0.125,0.125,0.125,0.125,0.125,0.125,0.125,0.125]
1	(0.1,0.9)	1.166	[0.057,0.330,0.262,0.125,0.057,0.057,0.057,0.057]
2	(0.5,0.5)	1.974	[0.035,0.521,0.162,0.141,0.035,0.035,0.035,0.035]
3	(0.2,0.8)	1.961	[0.017,0.256,0.385,0.069,0.221,0.017,0.017,0.017]

Source: Created by the authors.

Table 6. Predicted α, β, satisfaction and overall preference of each review.

t	$(\hat{\alpha}_t^J, \hat{\beta}_t^J)$	$\hat{S}_t^J$	$\hat{W}_t^J$
1	(0.5,0.5)	1.948	[0.077,0.269,0.221,0.125,0.077,0.077,0.077,0.077]
2	(0.3,0.7)	1.515	[0.030,0.563,0.141,0.144,0.030,0.030,0.030,0.030]
3	(0.367,0.633)	1.747	[0.020,0.297,0.351,0.080,0.192,0.020,0.020,0.020]
4	(0.3,0.7)	1.505	[0.009,0.475,0.291,0.037,0.16,0.009,0.009,0.009]

Source: Created by the authors.

Table 7. Product ranking for user J.

Rank	$\hat{E}_{5,PN}^J$	PN	Rank	$\hat{E}_{5,PN}^J$	PN	Rank	$\hat{E}_{5,PN}^J$	PN	Rank	$\hat{E}_{5,PN}^J$	PN
1	4.000	14	23	3.264	5	45	3.116	55	67	3.000	42
2	4.000	60	24	3.250	24	46	3.115	80	68	3.000	65
3	3.738	74	25	3.244	62	47	3.115	11	69	2.998	7
4	3.719	73	26	3.240	56	48	3.113	38	70	2.994	6
5	3.713	64	27	3.240	57	49	3.113	39	71	2.988	81
6	3.709	70	28	3.233	9	50	3.113	59	72	2.966	2
7	3.708	3	29	3.226	52	51	3.112	12	73	2.966	68
8	3.686	44	30	3.197	16	52	3.108	50	74	2.964	25
9	3.653	33	31	3.190	27	53	3.108	30	75	2.945	48
10	3.623	85	32	3.184	18	54	3.101	19	76	2.877	10
11	3.512	21	33	3.175	36	55	3.095	66	77	2.845	71
12	3.475	43	34	3.169	49	56	3.061	69	78	2.835	17
13	3.446	20	35	3.163	15	57	3.049	53	79	2.822	67
14	3.438	72	36	3.162	34	58	3.047	83	80	2.804	79
15	3.425	63	37	3.160	86	59	3.047	35	81	2.780	84
16	3.420	77	38	3.159	4	60	3.041	47	82	2.747	78
17	3.339	37	39	3.154	8	61	3.040	76	83	2.744	13
18	3.338	23	40	3.146	58	62	3.038	29	84	2.537	75
19	3.331	1	41	3.146	46	63	3.007	28	85	2.447	82
20	3.290	61	42	3.143	45	64	3.000	32	86	2.000	26
21	3.281	22	43	3.125	51	65	3.000	40			
22	3.281	31	44	3.116	54	66	3.000	41			

Note. PN = product number.
Source: Created by the authors.

preference of each review. For new consumers, we let both initial values of α and β be 0.5, and initial overall preference be [0.125,0.125,0.125,0.125,0.125,0.125,0.125, 0.125], which are shown in Table 5 when $t = 0$. Table 6 presents predicted values of α, β, the satisfaction and overall preference of each review. Every predicted overall preference is used to compute the expectation on all products of user J, which are regarded as the basis of product ranking.

As can be seen from Table 4, although user J has different sub-preferences in different time, it is obvious that he/she mainly focuses on four attributes: color, function, packaging and price. Correspondingly, the weights of color, function, packaging, and price in real and predicted overall preferences increase over several iterations. Because the number of reviews of user J is small, the values of the satisfaction parameters α and β fluctuate greatly. As the number of reviews increases, the satisfaction parameter will stabilise, which will help improve the accuracy of consumer preference prediction and product recommendation.

Table 7 presents the product ranking after the 4th review of user J. By using the predicted overall preference in Table 7, based on Equation (10), we compute the expectation on 86 products at the time of data collection. Then, we rank products according to expectations to make recommendations for user J. Based on online reviews, we calculate users' expectation on each product using the predicted overall consumer preference to rank products. But there is no guarantee that users will select the top-ranked products. As we collect more users' reviews, our preference prediction accuracy improves, and users are more likely to select the top-ranked products.

5. Conclusion

5.1. Theoretical implications

This paper considered an online retailer endowed with various products and the preferences of consumers on that platform are dynamic. The goal of this paper is to explore the efficient use of theories about consumer satisfaction formation for accurate dynamic preference elicitation from online reviews, so as to make personalised recommendations. To this end, this paper proposed a dynamic preference elicitation model based on the ECT where retailers use online reviews of customers as the source of preference elicitation. Theoretical contributions of this study are as follows:

First, in previous studies (Bhattacherjee, 2001), elements of ECT like expectation, perceived performance and satisfaction were usually collected on site with scales. Considering that it was not practical to collect the information of online consumers mentioned above frequently and repeatedly, this paper proposed formulas to define expectation, perceived performance and satisfaction, respectively. With the ECT as a framework, we extracted expectation, perceived performance and sub-preference from reviews of other consumers and the target customer, and then calculated confirmation and satisfaction based on the proposed formulas.

Second, different from the conventional ECT (Filieri et al., 2021; Lee & Kim, 2020), the influence of satisfaction on repurchase intention was replaced by the satisfaction that affects the influence degree of each sub-preference on the overall preference.

Third, existing studies about preference elicitation ignored the psychological perceptions and preference changes (Li et al., 2020; Shen et al., 2019). We considered the dynamics of the impact of expectation and confirmation on satisfaction in the ECT, and proposed a dynamic preference elicitation model. Based on the proposed model, we can get the expectation of a target customer for each product in which the overall preference was one of the factors and take the calculated expectation as the basis of product ranking and recommendation.

5.2. Managerial implications

This paper provided a method for retailers to elicit consumer preferences dynamically and recommend products based on online reviews. Managerial implications are as follows:

First, since the analysis objects are the detailed product attributes, managers can formulate corresponding strategies based on the analysis results. Taking the foundation data used in this paper as an example, there are eight attributes (*Brand*, *Color*, *Function*, *Packaging*, *Price*, *Quality*, *Smell*, *Other*) that production managers need to focus on to improve their products. In addition, marketing planners need to consider which product attributes should be highlighted in marketing.

Second, managers can obtain consumer preferences according to each users' review. After knowing the needs of consumers, online retailers can recommend products for consumers in line with their preferences, thus achieving the purpose of improving consumer satisfaction and repurchase intention.

5.3. Limitations and future research

Although this paper provided a method to help managers elicit consumer preferences, there were some limitations. First, we made assumptions to simplify the model, for instance, we used linear expressions of expectation and confirmation to represent satisfaction and assumed the ranges of parameters. In the application background of this paper, the variables of the simple ECT (Bhattacherjee, 2001) are the most necessary, so we use the simple ECT as the theoretical basis. In other application contexts, other variables besides the simple ECT variables may be required. For example, in expansions of the ECT (Shin et al., 2017), there were many other factors besides expectation and confirmation that affect satisfaction. Future studies can add other factors and consider other expressions of satisfaction based on this paper.

Second, we assumed a convex function relation between satisfaction and effect as shown in Figure 5 and applied it to the process of dynamic preference elicitation. The specific relation between satisfaction and effect needs to be further proved.

Finally, the limitations of the proposed method in practice are discussed. The proposed method can be seen as a small part of a recommendation system in practice, and its triggering should be conditional. Other programs of the recommendation system judge if customers have a high probability to buy a certain kind of product, and then the proposed model is run to calculate user expectations and other personal data. The triggering procedure of the proposed method is not within the scope of this study. The calculation challenge of this method lies in the need to calculate the trigger conditions, user expectations, and other personal data of numerous users at any time, which requires a huge server to ensure the smooth operation of the whole system. In addition, this paper presents a case study based on one product group, and further research is needed to see if the operation of the proposed method can be extended to other kinds of products.

Notes

1. https://www.emarketer.com/content/global-ecommerce-forecast-2021
2. Consumers may have consumed in other platforms or revealed personal characteristics in social media. However, owing to the confidentiality of data on every platform, it is supposed that sellers on different platforms can only infer expectations of consumers based on the information on their own online platforms.

3. Regular expression is a pattern used to match strings, whose functions include checking whether a string contains a certain substring, replacing the matched substring, and taking out the substring matching a certain condition from a string (Thompson, 1968).
4. http://stanfordnlp.github.io/CoreNLP/.
5. Considering that different products have different times to market and we cannot rank unsold items, we take the time of first review for each product as its beginning time to market. Products released later than the target customer's purchase time will not be considered.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

The work was supported by the National Natural Science Foundation of China (71771156, 71971145, 72171158), and the Sichuan Planning Project of Social Science (SC22A007).

ORCID

Keyu Lu http://orcid.org/0000-0003-3160-9452
Huchang Liao http://orcid.org/0000-0001-8278-3384

References

Bernstein, F., Modaresi, S., & Sauré, D. (2018). A dynamic clustering approach to data-driven assortment personalization. *Management Science*, *65*(5), 2095–2115. https://doi.org/10.1287/mnsc.2018.3031

Bhattacherjee, A. (2001). An empirical analysis of the antecedents of electronic commerce service continuance. *Decision Support Systems*, *32*(2), 201–214. https://doi.org/10.1016/S0167-9236(01)00111-7

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*, 993–1022. https://doi.org/10.1162/jmlr.2003.3.4-5.993

Dash, G., Kiefer, K., & Paul, J. (2021). Marketing-to-millennials: Marketing 4.0, customer satisfaction and purchase intention. *Journal of Business Research*, *122*, 608–620. https://doi.org/10.1016/j.jbusres.2020.10.016

eMarketer. (2021). Global ecommerce forecast 2021, eMarketer. Retrieved 11 November, 2011, from, https://www.emarketer.com/content/global-ecommerce-forecast-2021.

Fehr, E., & Hoff, K. (2011). Introduction: Tastes, castes and culture: The influence of society on preferences. *The Economic Journal*, *121*(556), F396–F412. https://doi.org/10.1111/j.1468-0297.2011.02478.x

Feng, H. Y., Tian, J., Wang, H. J. N., & Li, M. Q. (2015). Personalized recommendations based on time-weighted overlapping community detection. *Information & Management*, *52*(7), 789–800. https://doi.org/10.1016/j.im.2015.02.004

Filieri, R., Galati, F., & Raguseo, E. (2021). The impact of service attributes and category on eWOM helpfulness: An investigation of extremely negative and positive ratings using latent semantic analytics and regression analysis. *Computers in Human Behavior*, *114*, 106527. https://doi.org/10.1016/j.chb.2020.106527

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information Systems Research*, *19*(3), 291–313. https://doi.org/10.1287/isre.1080.0193

Fung, H. H., & Carstensen, L. L. (2003). Sending memorable messages to the old: Age differences in preferences and memory for advertisements. *Journal of Personality and Social Psychology*, *85*(1), 163–178. https://doi.org/10.1037/0022-3514.85.1.163
Giese, J. L., & Cote, J. A. (2000). Defining consumer satisfaction. *Academy of Marketing Science Review*, *1*(1), 1–22.
Godnov, U., & Redek, T. (2019). The use of user-generated content for business intelligence in tourism: Insights from an analysis of Croatian hotels. *Economic Research-Ekonomska Istraživanja*, *32*(1), 2455–2480. https://doi.org/10.1080/1331677X.2019.1633372
Han, H., & Hyun, S. S. (2018). Role of motivations for luxury cruise traveling, satisfaction, and involvement in building traveler loyalty. *International Journal of Hospitality Management*, *70*, 75–84. https://doi.org/10.1016/j.ijhm.2017.10.024
Helson, H. (1959). Adaptation-level theory. In: Sigmund Koch (ed.), *Psychology: A study of a science.* (Vol. 1, pp. 565–621). McGraw-Hill Book Company.
Hult, G. T. M., Sharma, P. N., Morgeson, F. V., III, & Zhang, Y. (2019). Antecedents and consequences of customer satisfaction: Do they differ across online and offline purchases? *Journal of Retailing*, *95*(1), 10–23. https://doi.org/10.1016/j.jretai.2018.10.003
Koren, Y. (2010). Collaborative filtering with temporal dynamics. *Communications of the ACM*, *53*(4), 89–97. https://doi.org/10.1145/1721654.1721677
Lee, J. H., & Kim, Y. K. (2020). Online reviews of restaurants: Expectation-confirmation theory. *Journal of Quality Assurance in Hospitality & Tourism*, *21*(5), 582–599. https://doi.org/10.1080/1528008X.2020.1712308
Liao, C. C., Lin, H. N., Luo, M. M. L., & Chea, S. (2017). Factors influencing online shoppers' repurchase intentions: The roles of satisfaction and regret. *Information & Management*, *54*(5), 651–668. https://doi.org/10.1016/j.im.2016.12.005
Li, H., Gupta, A., Zhang, J., & Flor, N. (2020). Who will use augmented reality? An integrated approach based on text analytics and field survey. *European Journal of Operational Research*, *281*(3), 502–516. https://doi.org/10.1016/j.ejor.2018.10.019
Li, W. H., Li, X. G., Deng, J. Z., Wang, Y., & Guo, J. P. (2021). Sentiment based multi-index integrated scoring method to improve the accuracy of recommender system. *Expert Systems with Applications*, *179*, 115105. https://doi.org/10.1016/j.eswa.2021.115105
Liu, Y. M., Wan, Y. Z., Shen, X. L., Ye, Z. Y., & Wen, J. (2021). Product customer satisfaction measurement based on multiple online consumer review features. *Information*, *12*(6), 234. https://doi.org/10.3390/info12060234
Mauri, A. G., & Minazzi, R. (2013). Web reviews influence on expectations and purchasing intentions of hotel potential customers. *International Journal of Hospitality Management*, *34*, 99–107. https://doi.org/10.1016/j.ijhm.2013.02.012
Mittal, V., & Kamakura, W. A. (2001). Satisfaction, repurchase intent, and repurchase behavior: Investigating the moderating effect of customer characteristics. *Journal of Marketing Research*, *38*(1), 131–142. https://doi.org/10.1509/jmkr.38.1.131.18832
Montejo-Ráez, A., Martínez-Cámara, E., Martín-Valdivia, M. T., & Ureña-López, L. A. (2014). Ranked wordnet graph for sentiment polarity classification in twitter. *Computer Speech & Language*, *28*(1), 93–107. https://doi.org/10.1016/j.csl.2013.04.001
Oliver, R. L. (1980). A cognitive model of the antecedents and consequences of satisfaction decisions. *Journal of Marketing Research*, *17*(4), 460–469. https://doi.org/10.2307/3150499
Oliver, R. L. (2014). *Satisfaction: A Behavioral Perspective on the Consumer*. Routledge.
Parasuraman, A., Zeithaml, V. A., & Berry, L. (1988). SERVQUAL: A multiple-item scale for measuring consumer perceptions of service quality. *Journal of Retailing*, *64*(1), 12–40.
Park, J. H., & Lee, B. K. (2021). An opinion-driven decision-support framework for benchmarking hotel service. *Omega*, *103*, 102415. https://doi.org/10.1016/j.omega.2021.102415
Park, S., & Nicolau, J. L. (2015). Asymmetric effects of online consumer reviews. *Annals of Tourism Research*, *50*, 67–83. https://doi.org/10.1016/j.annals.2014.10.007
Pavlou, P. A., & Dimoka, A. (2006). The nature and role of feedback text comments in online marketplaces: Implications for trust building, price premiums, and seller differentiation. *Information Systems Research*, *17*(4), 392–414. https://doi.org/10.1287/isre.1060.0106

Pereira, F. S., Gama, J., de Amo, S., & Oliveira, G. M. (2018). On analyzing user preference dynamics with temporal social networks. *Machine Learning*, *107*(11), 1745–1773. https://doi.org/10.1007/s10994-018-5740-2

Prayag, G., Hassibi, S., & Nunkoo, R. (2019). A systematic review of consumer satisfaction studies in hospitality journals: conceptual development, research approaches and future prospects. *Journal of Hospitality Marketing & Management*, *28*(1), 51–80. https://doi.org/10.1080/19368623.2018.1504367

Rafeh, R., & Bahrehmand, A. (2012). An adaptive approach to dealing with unstable behaviour of users in collaborative filtering systems. *Journal of Information Science*, *38*(3), 205–221. https://doi.org/10.1177/0165551512437517

Sahoo, N., Singh, P. V., & Mukhopadhyay, T. (2012). A hidden Markov model for collaborative filtering. *MIS Quarterly*, *36*(4), 1329–1356. https://doi.org/10.2307/41703509

Shen, R. P., Zhang, H. R., Yu, H., & Min, F. (2019). Sentiment based matrix factorization with reliability for recommendation. *Expert Systems with Applications*, *135*, 249–258. https://doi.org/10.1016/j.eswa.2019.06.001

Shi, Y. C. (2014, February 16–19). *An improved collaborative filtering recommendation method based on timestamp* [Paper presentation]. 16th International Conference on Advanced Communication Technology, PyeongChang, South Korea. IEEE (pp. 784–788).

Shin, D. H., Lee, S. J., & Hwang, Y. J. (2017). How do credibility and utility play in the user experience of health informatics services? *Computers in Human Behavior*, *67*, 292–302. https://doi.org/10.1016/j.chb.2016.11.007

Soumya, S., & Pramod, K. V. (2020). Sentiment analysis of Malayalam tweets using machine learning techniques. *ICT Express*, *6*(4), 300–305. https://doi.org/10.1016/j.icte.2020.04.003

Sparks, B. A., So, K. K. F., & Bradley, G. L. (2016). Responding to negative online reviews: The effects of hotel responses on customer inferences of trust and concern. *Tourism Management*, *53*, 74–85. https://doi.org/10.1016/j.tourman.2015.09.011

Štefko, R., Bačík, R., Fedorko, R., Oleárová, M., & Rigelský, M. (2019). Analysis of consumer preferences related to the use of digital devices in the e-commerce dimension. *Entrepreneurship and Sustainability Issues*, *7*(1), 25–33. https://doi.org/10.9770/jesi.2019.7.1(2)

Thompson, K. (1968). Regular search expression algorithm. *Communications of the ACM*, *11*(6), 419–422. https://doi.org/10.1145/363347.363387

Yuksel, A., & Yuksel, F. (2001). Measurement and management issues in customer satisfaction research: Review, critique and research agenda: Part one. *Journal of Travel & Tourism Marketing*, *10*(4), 47–80. https://doi.org/10.1300/J073v10n04_03

Zhang, C. X., Xu, Z. S., Gou, X. J., & Chen, S. X. (2021). An online reviews-driven method for the prioritization of improvements in hotel services. *Tourism Management*, *87*, 104382. https://doi.org/10.1016/j.tourman.2021.104382

Zhang, K. Z., Zhao, S. J., Cheung, C. M., & Lee, M. K. (2014). Examining the influence of online reviews on consumers' decision-making: A heuristic–systematic model. *Decision Support Systems*, *67*, 78–89. https://doi.org/10.1016/j.dss.2014.08.005

Zhao, W., Guan, Z., Chen, L., He, X. F., Cai, D., Wang, B. D., & Wang, Q. (2018). Weakly-supervised deep embedding for product review sentiment analysis. *IEEE Transactions on Knowledge and Data Engineering*, *30*(1), 185–197. https://doi.org/10.1109/TKDE.2017.2756658

Zheng, N., & Li, Q. D. (2011). A recommender system based on tag and time information for social tagging systems. *Expert Systems with Applications*, *38*(4), 4575–4587. https://doi.org/10.1016/j.eswa.2010.09.131