

Crowd intelligence: Analyzing online product reviews for preference measurement



Shengsheng Xiao^a, Chih-Ping Wei^{b,*}, Ming Dong^a

^a Department of Operations Management, Antai College of Economics & Management, Shanghai Jiao Tong University, Shanghai 200052, China

^b Department of Information Management, National Taiwan University, Taipei 10617, Taiwan

ARTICLE INFO

Article history:

Received 21 October 2014

Received in revised form 11 August 2015

Accepted 29 September 2015

Available online 8 October 2015

Keywords:

Preference analysis and measurement

Aggregate consumer preferences

Crowd intelligence

Online product reviews

User generated content

Kano model

ABSTRACT

The proliferation of product review websites produces a large, publicly accessible information resource for firms that seek to understand consumers' preferences. To facilitate product design or improvement, we propose a novel econometric preference measurement model, the modified ordered choice model (MOCM), to extract aggregate consumer preferences from online product reviews. Moreover, to categorize customer requirements on the basis of the aggregate consumer preferences estimated by the MOCM model, we extend the Kano model and propose a marginal effect-based Kano model (MEKM). We empirically evaluate the effectiveness of the proposed MOCM model and demonstrate the utility of the proposed MEKM model.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

With the advances and rapid proliferation of Web 2.0 innovations, people increasingly use various online word-of-mouth (WOM) channels to share their consumption experiences with and preferences for a wide range of products. The resulting collection of online product reviews has become an important information source that consumers use to decide whether to purchase a product or which product to select. According to one recent report,¹ 77% of consumers read online product reviews before shopping, and 75% trust online product reviews more than personal recommendations. In addition, 81% of users indicate that they receive helpful advice from online product reviews. Online product reviews thus constitute a novel type of WOM online (commonly referred to as eWOM) [11,23,38,62].

Compared with traditional WOM, online product reviews provide a more publicly accessible information source to understand consumer perceptions and preferences; traditionally, such data were difficult to collect on a large scale in the offline world. Because of this characteristic, substantial literature seeks to connect some measurable attributes of online product reviews

(e.g., review valence, volume, product rating, product comparison) with consumers' purchase behaviors [38,48], product sales [9,17,34,61], firms' economic outcomes [3,19] and operational strategies [1,9,10,34,25,47,57]. Some prior studies also attempt to exploit this public information source and turn individual consumer opinions into aggregate consumer preference measures (i.e., a preference measurement model) that capture the effects of consumers' sentiment toward product features on product ratings [13,32,37,40].

Measuring such aggregate consumer preferences is critical to firms, because it can facilitate their planning and decision making pertaining to product improvement, new product development, pricing, market segmentation, positioning, and advertising [18,36]. Conjoint analysis has been the main quantitative preference measurement method since it was first introduced by Green and Rao [20]. As a result, preference measurement is often equated with conjoint analysis [46]. To conduct conjoint analysis, we need to collect consumers' preferences through surveys or experiments, which require rigorous design and involve a proper procedure to ensure the quality of responses from respondents. There is no doubt that such data collection approach is time consuming and costly. In contrast, the availability of large-scale online product reviews offers the promise of an alternative means to measure aggregate consumer preferences. Compared with the traditional data collection approach for conjoint analysis, using online product reviews to support preference measurement has several advantages. First, online product reviews are publicly

* Corresponding author. Tel.: +886 2 3366 1181; fax: +886 2 3366 9455.

E-mail addresses: xiaoshengsheng@sjtu.edu.cn (S. Xiao), cpwei@ntu.edu.tw (C.-P. Wei), mdong@sjtu.edu.cn (M. Dong).

¹ <http://wallblog.co.uk/2012/05/14/the-future-of-the-social-customer-rise-of-fan-commerce-infographic/>.

available and can be collected easily. Second, online product reviews are voluntarily produced by actual consumers [13,46] and do not depend on surveys or respondents. Prior studies have shown that consumer opinions expressed in online product reviews offer a good proxy for the overall WOM of the products being discussed and then become a new source of preference data [3,13,40,62]. Thus, online product reviews represent a more representative preference dataset than those collected by surveys or experiments. Third, the size of online product reviews is generally large and they often cover diverse product features. Consequently, these online product reviews can be employed to construct more comprehensive preference measurement models than preference datasets collected through surveys or experiments. Considering these advantages, we focus on the use of online product reviews for preference measurement and investigate two pertinent research questions: (1) How can we effectively measure aggregate consumer preferences from online product reviews? and (2) How can we categorize customer requirements, on the basis of the estimated aggregate consumer preferences?

To address these research questions, we first establish a framework to extract, from online product reviews, the product features discussed and the reviewers' sentiment orientations (like or dislike) toward them. To better explain reviewers' rating behavior, we also collect additional information about the reviewers, including the total number of reviews written by each of them and the trust relations among them, for calibrating our preference measurement model. Next, we propose an econometric model, referred to as the modified ordered choice model (MOCM), to measure aggregate consumer preferences from online product reviews. This new model takes into consideration the heteroscedasticity of reviewers' rating variance and allows reviewers to assign rating scores according to their own thresholds. Furthermore, to categorize customer requirements and support product design on the basis of the aggregate consumer preferences estimated by our proposed MOCM model, we extend the Kano model [8,28,30,56] and propose a marginal effect-based Kano model (MEKM). Finally, we empirically evaluate the effectiveness of our econometric preference measurement model (i.e., MOCM) and demonstrate the utility of our proposed MEKM model. Using a dataset collected from Epinions.com, we show that our proposed MOCM model outperforms existing models and the MEKM model provides a viable method for further categorizing and prioritizing customer requirements.

The remainder of the paper is organized as follows. Section 2 reviews the literature relevant to this study. We describe our preprocessing procedure for extracting product features and reviewers' sentiment orientations from online product reviews and then depict our data coding scheme in Section 3. In Section 4, we detail our proposed econometric model (i.e., MOCM) for estimating aggregate consumer preferences from online product reviews. Section 5 depicts our proposed MEKM model for categorizing customer requirements, based on the aggregate consumer preferences estimated by MOCM. In Section 6, we report our empirical study, in which we use online product reviews from Epinions.com to evaluate the effectiveness of our proposed MOCM model and illustrate the utility of the proposed MEKM model. We conclude in Section 7 by highlighting our theoretical and practical contributions as well as some further research directions.

2. Literature review

In this section, we review existing preference measurement methods and analyze their limitations to justify our research motivation. In addition, we summarize the Kano model, which provides the basis for our MEKM model.

2.1. Existing preference measurement methods

In the past several decades, many methods have been developed to measure aggregate consumer preferences quantitatively. Depending on the data they use, existing preference measurement methods can be classified into three major approaches: survey-, behavior-, and online review-based.

Conjoint analysis was first introduced by Green and Rao [20]. Since its introduction, preference measurement is often linked to conjoint analysis [46]. One significant characteristic of conjoint analysis is that it depends strongly on survey data, collected through surveys or experiments. As a result, the survey-based approach mainly refers to those preference measurement methods that use conjoint analysis or its variants. Using consumers' preference data collected from surveys or experiments, the survey-based approach typically relies on econometric and statistical methods to analyze these data and determine how people value the different features that constitute an individual product or service [24,52].

The survey-based approach is formal and rigorous, but its data collection process is time consuming and costly. To overcome these challenges, some studies exploit the use of consumers' behavioral data, collected from shopping environments, to infer aggregate consumer preferences. Different from the survey-based approach, the behavior-based approach uses the data about consumers' behavior directly (e.g., items placed in shopping carts, items purchased, shopping paths), rather than preference data collected from surveys and experiments. For example, with point-of-sales data, Fader and Hardie [16] employ a discrete choice model to measure consumer preferences for selected product features. In contrast, Hui et al. [27] estimate aggregate consumer preferences from consumers' purchases (i.e., transaction data) and their shopping paths (collected by RFID technology). Because the data about consumers' behavior may have been collected already by retailers (e.g., consumers' purchases) or can be acquired easily (e.g., with the support of some information technologies), the data collection cost of this approach tends to be far less than that of the survey-based approach.

Finally, the increasing availability and accessibility of online product reviews contributed by consumers have prompted some studies to investigate ways to measure aggregate consumer preferences from online product reviews. This online review-based approach leverages the large collection of existing, publicly available online product reviews and thus represents an appealing alternative. Archak et al. [4] propose a hedonic regression approach to analyze the strength and polarity of consumer review opinions. However, they did not consider opinion heterogeneity expressed in each review. In a follow-up study, they note that consumers' preferences can be reflected by online product reviews but do not propose a detailed method to extract them [18]. Lee and Bradlow [36] emphasize the importance of online product reviews for conjoint analyses in marketing and propose a text mining technique to extract the product features discussed in online product reviews, as well as consumers' sentiment orientations toward these features. Li et al. [40] develop a social intelligence mechanism to extract and consolidate the reviews expressed via social media and to derive insights to help firms make decisions on product portfolio design. Decker and Trusov [13] propose three econometric models (i.e., Poisson regression, negative binomial regression, and latent class Poisson regression models) to measure aggregate consumer preferences from online product reviews about mobile phones.

Several online review-based preference measurement models have been proposed [13,36,37]. However, they incur several limitations. First, it may be not suitable to use counting models, such as Poisson regression, negative binomial regression, and latent class Poisson regression models, to assess reviewers' rating

behavior (i.e., measure aggregate consumer preferences). That is mainly because the ratings in online product reviews use an ordinal rather than interval scale, and counting models are not designed to model ordinal scale data. Second, existing online review-based preference measurement methods consider review content only (i.e., textual comments and rating scores), without acknowledging that different reviewers have their own biases when giving rating scores, and their rating behaviors may be influenced by their personal characteristics [39,53]. Therefore, to better explain reviewers' rating behavior and improve the effectiveness of preference measurement, preference measurement models should take reviewer characteristics into consideration [13].

Following the stream of preference measurement literature, we propose in this study a modified ordered choice model (MOCM) to extract aggregate consumer preferences from online product reviews. Moreover, we incorporate some covariates from reviewer profiles into the proposed model, to improve the effectiveness of preference measurement.

2.2. Kano model

Our research also relates to the Kano model, a two-dimensional model first developed by Kano et al. [30]. The Kano model classifies product features (attributes) into different categories (see Fig. 1), according to the degree of fulfillment of the product feature and its effect on customer satisfaction [8,28,30,56]. Specifically,

1. *Must-be (or basic) features*: These product features are taken for granted when fulfilled, but if they are not fulfilled, customers become dissatisfied.
2. *Performance (or one-dimensional) features*: A product feature is considered as a performance feature if its fulfillment is positively associated with customer satisfaction.
3. *Excitement (or attractive) features*: Excitement features are the opposites of must-be features. That is, excitement features offer satisfaction when fulfilled but do not result in dissatisfaction if not fulfilled.
4. *Indifferent features*: When the degree of fulfillment of a product feature is either not associated or only marginally associated with customer satisfaction or customer dissatisfaction, this feature is referred to as an indifferent feature.
5. *Reverse features*: Reverse features imply that when the degree of fulfillment increases, customers become more dissatisfied.

The Kano model is typically constructed on the basis of customer surveys, consisting of a set of question pairs. Each

question pair pertains to a product feature and includes a functional form question, which captures customer responses if the product has this feature, and a dysfunctional form question to capture customer responses if the product does not have this feature. Customers' responses to the question pair reveal their perceptions about that corresponding feature. The questionnaire is deployed to a number of customers, and the answers to each question pair is aligned with the Kano evaluation table [5]. The classification of product features is then made on the basis of a statistical analysis of the survey results of all respondents.

As mentioned, the Kano model involves two dimensions: the degree of fulfillment of product features and the effect on customer satisfaction. However, aggregate consumer preferences measured by an online review-based preference measurement method cannot map directly to these two dimensions. Hence, in this study, we extend the Kano model by employing marginal effect information disclosed by the proposed MOCM model. Specifically, we follow the feature classification framework of the Kano model but use the marginal effect information derived from the MOCM model to quantify consumers' satisfaction with different product features. To the best of our knowledge, this study is the first to combine the concept of the Kano model with a preference measurement model and thereby categorize product features.

3. Data preprocessing framework

In this section, we first depict the preprocessing procedure that we use to extract product features and reviewers' sentiment orientations from online product reviews. We then discuss additional data (i.e., reviewer profiles) that our proposed preference measurement model requires. Finally, we describe our data coding scheme in detail.

3.1. Product feature extraction from online product reviews

To understand why reviewers assign different rating scores to products, we should first extract product features discussed in these reviews and the reviewers' sentiment orientations (positive or negative) toward these product features. Prior studies have developed various opinion mining methods for this purpose using natural language processing or text mining techniques [7,12,26,49,55,58,59]. Because we aim to measure aggregate consumer preferences from online product reviews, developing a new product feature extraction method is not the main focus of our study; instead, we employ and extend existing methods to support our product feature extraction. Moreover, this study takes semi-structured (or even structured) reviews,² rather than free-text reviews, as inputs due to the following reasons. First, consumers explicitly express their sentiment orientations toward product features in semi-structured reviews (i.e., pro and con phrases are organized separately and directly convey reviewers' sentiment orientations toward product features). As a result, given an online product review, we do not need to analyze the full text of the review to identify the sentiment orientations of the reviewer toward different product features. This will greatly improve the effectiveness and efficiency of the data preprocessing stage. Second, many studies reveal that the pro/con phrases in semi-structured reviews generally summarize the opinions expressed in the full text of the corresponding reviews [6,31]. Therefore, using

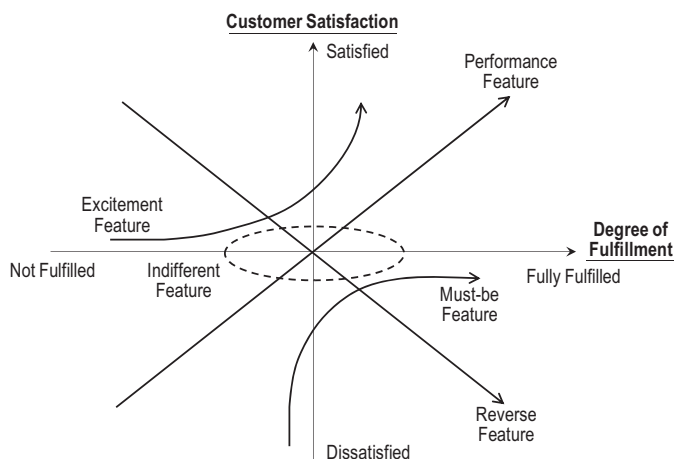


Fig. 1. Classification of features in the Kano model.

² Currently, three review formats predominate online: (1) structured reviews, in which reviewers identify the pros and cons of products and put them in designated fields; (2) semi-structured reviews that allow reviewers to complete the designated fields for pros and cons but also write detailed reviews in free-text format; and (3) free-text reviews, in which reviewers only write detailed reviews in free-text format.

these semi-structured reviews should not lead to the loss of too much useful information, compared with using free-text reviews. Third, a semi-structured review is a popular format used by many online product review platforms (e.g., Epinions.com, 360buy.com, Taobao.com). Thus, the product feature extraction procedure developed and employed in this study can be applied to online product reviews collected from various platforms.

With a collection of semi-structured product reviews, our product feature extraction procedure consists of five steps (left column, Fig. 2):

- (1) Partitioning, stemming and removing stop words. We start by splitting the pros and cons in each review into individual phrases, then reduce the inflected (or sometimes derived) words to their stems, bases, or root forms. Finally, we remove stop words and other meaningless words (e.g., “the,” “an,” “nothing else”) from the partitioned and stemmed pro/con phrases.
- (2) Extracting explicit features. Explicit product features can be found easily in the pro/con fields, but they often are accompanied by positive or negative adjectives that reiterate the reviewers’ sentiment orientations. Because we already know the reviewer’s sentiment orientation about a product feature, depending on whether it appears in the pro or con field, such positive or negative adjectives are not needed in subsequent analyses. Thus, we remove these adjectives to obtain explicit product features in this step. For example, reviewers may put “good touch screen” and “nice camera” in the pro field. In this case, “good” and “nice” are positive adjectives that recap the reviewers’ positive attitude toward these two product features. Accordingly, we remove “good” and “nice” and reduce “good touch screen” and “nice camera” to “touch screen” and “camera,” respectively.
- (3) Transforming implicit features. Reviewers might implicitly include product features in the pro/con fields. For example, “expensive” is an implicit product feature, corresponding to the explicit product feature “price.” In this step, we identify and transform all implicit features into explicit features. To ensure its accuracy, we perform this task manually.

- (4) Merging and grouping product features. People often use different terms to describe the same concept (or object) [54]. To address this word mismatch problem, we merge terms in different reviews that refer to the same product feature and replace them with a standardized term for this feature. We employ a semi-automated method for the described merging and grouping task. Specifically, we first merge product features with a text mining package (i.e., IBM SPSS Text Analytics). Built on thousands of commonly used words and phrases, this package automatically calculates the similarity between product features and merges them according to their similarities. Subsequently, we manually check the grouping results and make adjustments if needed.
- (5) Pruning infrequent product features. This step removes product features that are less frequent in the target collection of product reviews. For this study, we set a threshold of 10%; that is, we remove product features that occur in less than 10% of product reviews.

After the final step, we retain a list of extracted product features for each semi-structured product review in the target collection. An illustration of our product feature extraction procedure is in the right column of Fig. 2. It is worth noting that, although we take semi-structured reviews as inputs, our study can easily be extended to handle free-text reviews. In this case, we can employ an existing opinion mining method for extracting product features from free-text reviews and identifying reviewers’ sentiment orientations toward these product features. Subsequently, we perform implicit feature transformation (step 3 in Fig. 2), product feature grouping (step 4), and infrequent feature pruning (step 5) and thus conclude the product feature extraction procedure for free-text reviews.

3.2. Collection of reviewer profiles for preference measurement purposes

In addition to extracting product features and identifying reviewers’ sentiment orientations from online product reviews, we also need to collect reviewer profiles, which help us to construct

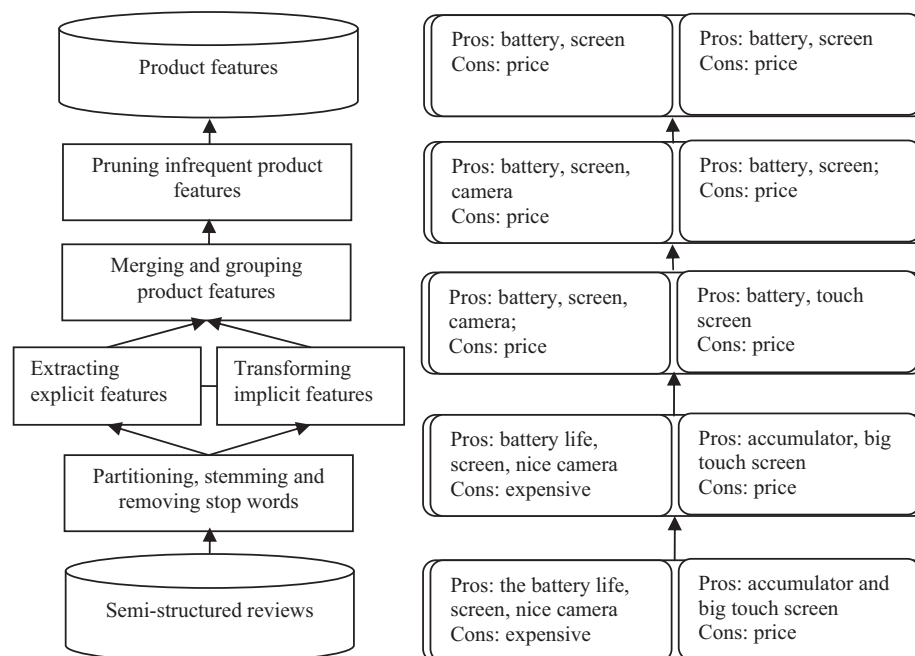


Fig. 2. Procedure of product feature extraction and an example. Note: “Camera” is assumed to be an infrequent product feature in this example.

more covariates to explain reviewers' rating behavior and improve the effectiveness of preference measurement. In a typical online product review platform, a reviewer profile contains the reviewer's basic information (e.g., geographic and demographic information, nickname) and records his/her historical activities on this platform (e.g., total number of reviews written, trust relations with other reviewers). We are mainly interested in two types of information. First, we consider the total number of reviews written by each reviewer; compared with reviewers who have written fewer reviews, those who have published more usually are more experienced with expressing their preferences. Thus, the rating behavior of reviewers who have written more reviews should be different from that of reviewers who have written fewer reviews. Second, trust relations between reviewers likely disclose each reviewer's overall position or reputation in this community. For our empirical study, we construct reviewers' trust network and use the normalized PageRank value [35,43] of each reviewer to signify his or her overall position or reputation.

3.3. Data coding scheme

After our data preprocessing, we use the following scheme for data coding. Specifically, we use subscripts $k = 1, 2, 3, \dots, K$ to index individual reviews; $l = 1, 2, 3, \dots, L$ to denote product features mentioned in the reviews; $i = 1, 2, 3, \dots, I$ to represent reviewers; and $s \in \{\text{pro}, \text{con}\}$ to represent the sentiment orientation of each extracted product feature in a review (whether a feature is a pro or con in a review). We let $x_{k,l,\text{pro}} = 1$, if product feature l is taken as a pro in review k (product feature l appears in the pro field of review k), and $x_{k,l,\text{pro}} = 0$ otherwise. Similarly, we let $x_{k,l,\text{con}} = 1$ if product feature l is taken as a con in review k (product feature l appears in the con field of review k), and $x_{k,l,\text{con}} = 0$ otherwise. Product brand mentioned in review k is represented by B_k (B_k is actually a dummy variable that serves as the brand-specific intercept and is incorporated to control brand effect), the total number of reviews written by reviewer i is represented by n_i , and the overall reputation of reviewer i (measured by the normalized PageRank value) is denoted by tr_i .

4. MOCM: a new preference measurement model

Considering the ordinal scale of the rating scores in online product reviews, we do not intend to use the counting models employed by Decker and Trusov [13] to explain reviewers' rating behavior (instead, they serve as benchmark models for our empirical study). In this section, we first explain how the basic ordered choice model (BOCM) measures aggregate consumer preferences. Subsequently, we relax two assumptions of the BOCM model by taking the heterogeneity of each reviewer's rating behavior into consideration to establish the modified ordered choice model (MOCM).

4.1. Basic ordered choice model (BOCM)

The BOCM model is proposed by McKelvey and Zavoina [42] for the analysis of categorical, non-quantitative choices, outcomes, and responses. Because of its capability to deal with ordinal data, we use this model as a basis to develop our preference measurement model. Specifically, the BOCM model assumes that the rating score given in online product review k is based on the following latent regression:

$$y_k^* = \beta' \cdot (\mathbf{x}_{k,\text{pro}}, \mathbf{x}_{k,\text{con}}, B_k) + \varepsilon_k \quad (1)$$

where ε_k is a random variable following a normal distribution. The continued latent variable, y_k^* , is observed in discrete form through a

censoring mechanism:

$$y_k = \begin{cases} 0 & \text{if } -\infty < y_k^* \leq \mu_0 \\ 1 & \text{if } \mu_0 < y_k^* \leq \mu_1 \\ 2 & \text{if } \mu_1 < y_k^* \leq \mu_2 \\ \dots & \\ j & \text{if } \mu_{j-1} < y_k^* \leq +\infty \end{cases} \quad (2)$$

$\mathbf{x}_{k,\text{pro}}$ and $\mathbf{x}_{k,\text{con}}$ are vectors containing dummy variables. If we assume review k is written by reviewer i , then $\mathbf{x}_{k,\text{pro}}$ and $\mathbf{x}_{k,\text{con}}$ indicate the sentiment orientations of reviewer i toward all the extracted product features. More specifically, $x_{k,l,\text{pro}} = 1$ (where $x_{k,l,\text{pro}} \in \mathbf{x}_{k,\text{pro}}$) means that product feature l in review k is regarded as a "pro" by reviewer i and $x_{k,l,\text{pro}} = 0$ denotes that reviewer i does not consider product feature l in review k a "pro." The meaning of $x_{k,l,\text{con}}$ is similar to $x_{k,l,\text{pro}}$. In addition, B_k represents the brand name of the focal product discussed in review k . The outcome y_k refers to the observable rating score given in review k . The unknown marginal utilities β , and unknown thresholds in Eq. (2) must be estimated, using observable online product review samples (indexed by $k = 1, \dots, K$). In the following section, we use $\mathbf{x}_{k,\cdot}$ to denote $\mathbf{x}_{k,\text{pro}}$ and $\mathbf{x}_{k,\text{con}}$. The estimates in this model can be obtained with a maximum likelihood method. In the context of the BOCM model, the probability entering the log likelihood function is by Eq. (3) (where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the normal distribution):

$$\begin{aligned} \text{Prob}[y_k = j | \mathbf{x}_{k,\cdot}, B_k] &= \begin{cases} \Phi[\mu_0 - \beta' \cdot (\mathbf{x}_{k,\cdot}, B_k)], & \text{when } j = 0 \\ \Phi[\mu_j - \beta' \cdot (\mathbf{x}_{k,\cdot}, B_k)] - \Phi[\mu_{j-1} - \beta' \cdot (\mathbf{x}_{k,\cdot}, B_k)], & \text{when } 0 < j < J \\ 1 - \Phi[\mu_{j-1} - \beta' \cdot (\mathbf{x}_{k,\cdot}, B_k)], & \text{when } j = J \end{cases} \end{aligned} \quad (3)$$

4.2. Development of the modified ordered choice model (MOCM)

Although it is convenient to obtain estimates from the BOCM model, an issue for the estimation is the unobserved heterogeneity of each reviewer. Reviewers differ from one another, such that they express different opinions about the same products and assign different rating scores. If we reexamine the BOCM model, we will find two sources of individual heterogeneity: the random component ε_i and the thresholds μ_j , used by each reviewer. In the following sections, we will discuss how to modify the BOCM model by taking these two heterogeneity patterns into consideration.

4.2.1. Heteroscedasticity of rating behavior

In the BOCM model, $\text{Var}[\varepsilon_k | \mathbf{x}_{k,\cdot}, B_k] = 1$, so $\text{Var}[y_k | \mathbf{x}_{k,\cdot}, B_k] = 1$. It means that the rating variance is identical across different reviewers. Obviously, this is not a reasonable assumption. Intuitively, a reviewer who has written many reviews should be better expressing his or her opinions and assigning rating scores that are closer to products' true ratings than a reviewer who has published only a few reviews. The rating variance from a more experienced reviewer then should differ from that of a less experienced reviewer. We first relax this homogeneous variance assumption by letting $\text{Var}[\varepsilon_k | \mathbf{x}_{k,\cdot}, B_k]$ satisfy Eq. (4) (assuming review k is written by reviewer i).

$$\text{Var}[\varepsilon_k | \mathbf{x}_{k,\cdot}, B_k, \mathbf{w}_i] = (\sigma_k^i)^2, \quad \text{where } (\sigma_k^i)^2 = \exp(\gamma' \mathbf{w}_i + \nu) \quad (4)$$

where \mathbf{w}_i is a vector of variables to explain reviewer i 's rating variance (in our empirical study, we use the total number of reviews written by reviewer i to represent \mathbf{w}_i), and γ is the associated parameter set. Furthermore, ν ($\nu \sim N[0, 1]$) is a standard latent random variable for interpreting the unobservable part of

the error term for rating variance. We use the exponential function, as widely used in prior literature (e.g., [15,21]), to ensure the positivity of the parameterized σ_k^i . After considering heteroscedasticity, we can replace the conditional probability entering the log likelihood function by Eq. (5):

$$\text{Prob}[y_i = j | \mathbf{x}_{k,i}, B_k, \mathbf{w}_i, v] = \begin{cases} \Phi \left[\frac{\mu_{i,0} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right], & \text{if } j = 0 \\ \frac{\Phi \left[\frac{\mu_{i,j} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right]}{\Phi \left[\frac{\mu_{i,j} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right]} - \Phi \left[\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right], & \text{if } 0 < j < J \\ 1 - \Phi \left[\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right], & \text{if } j = J \end{cases} \quad (5)$$

4.2.2. Different thresholds for different reviewers

Another assumption in the BOCM model pertains to thresholds μ_j ($j = 0, 1, 2, \dots, J-1$), assumed to be identical across different reviewers. In reality, different reviewers assign the rating scores, according to their own standards. As a result, these thresholds should take different values for different reviewers. We therefore model the reviewer-specific thresholds (Eluru et al. [15] and Greene and Hensher [21] use similar methods to solve different problems), as follows:

$$\mu_{i,j} = \mu_{i,j-1} + e^{\theta_j + \delta' \mathbf{z}_i + \eta} \quad (j = 0, 1, 2, \dots, J-1) \quad (6)$$

The continuous latent variable y_k^* is the observed rating score in discrete form, through a censoring mechanism:

$$y_k = \begin{cases} 0 & \text{if } -\infty < y_k^* \leq \mu_{i,0} \\ 1 & \text{if } \mu_{i,0} < y_k^* \leq \mu_{i,1} \\ 2 & \text{if } \mu_{i,1} < y_k^* \leq \mu_{i,2} \\ \dots & \dots \\ J & \text{if } \mu_{i,J-1} < y_k^* \leq +\infty \end{cases} \quad (8)$$

In Eqs. (7) and (8), $\sigma_k^i = \exp(\gamma' \mathbf{w}_i + v)$ and $\mu_{i,j} = -\mu_{i,j-1} + \exp(\theta_j + \delta' \mathbf{z}_i + \eta)$. Estimates in the MOCM model can be obtained by the maximum simulated likelihood method. Now, the modified conditional probability entering the log likelihood function is replaced by:

$$\text{Prob}[y_k = j | \mathbf{x}_{k,i}, B_k, \mathbf{w}_i, \mathbf{z}_i, v, \eta] = \begin{cases} \Phi \left[\frac{\mu_{i,0} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right], & \text{if } j = 0 \\ \frac{\Phi \left[\frac{\mu_{i,j} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right]}{\Phi \left[\frac{\mu_{i,j} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right]} - \Phi \left[\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right], & \text{if } 0 < j < J \\ 1 - \Phi \left[\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right], & \text{if } j = J \end{cases} \quad (9)$$

Because the term used in the log likelihood function is conditional on the unobservable variables v and η , the unconditional probability of Eq. (9) can be written further as:

$$\text{Prob}[y_k = j | \mathbf{x}_{k,i}, B_k, \mathbf{w}_i, \mathbf{z}_i] = \begin{cases} \int_{v,\eta} \left\{ \Phi \left[\frac{\mu_{i,0} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right] \right\} f(v) f(\eta) dv d\eta, & \text{if } j = 0 \\ \int_{v,\eta} \left\{ \Phi \left[\frac{\mu_{i,j} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right] - \Phi \left[\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right] \right\} f(v) f(\eta) dv d\eta, & \text{if } 0 < j < J \\ 1 - \int_{v,\eta} \left\{ \Phi \left[\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right] \right\} f(v) f(\eta) dv d\eta & \text{if } j = J \end{cases} \quad (10)$$

As Eq. (6) reveals, the thresholds used by different reviewers are no longer the same but instead depend on θ_j (specific to the j -star rating score, but independent to reviewers), a parameter vector \mathbf{z}_i , and a latent variable η . Moreover, θ_j illustrates the fixed effect of a reviewer giving a j -star rating score ($j = 0, 1, 2, \dots, J$). The variable vector \mathbf{z}_i contains variables to capture reviewer-specific choices when they assign rating scores (in our empirical study, we use the overall reputation score of reviewer i , which is measured by the normalized PageRank value according to the reviewers' trust network, to represent \mathbf{z}_i); η follows a normal distribution and serves to capture unobservable bias. It is worth noting that η and v in Eqs. (5) and (6) are independent. By using Eq. (6), we ensure that $\mu_{i,j}$ is always larger than $\mu_{i,j-1}$.

4.2.3. Our proposed modified ordered choice model (MOCM)

By combining the aforementioned modifications, we establish the modified ordered choice model (MOCM) to explain reviewers' rating behaviors. If we assume review k is posted by reviewer i , this model can be written as:

$$y_k^* = \beta' \cdot (\mathbf{x}_{k,i}, B_k) + \varepsilon_k, \quad \text{where } \text{Var}[\varepsilon_k | \mathbf{x}_{k,i}, B_k, \mathbf{w}_i] = (\sigma_k^i)^2 \quad (7)$$

It is difficult to calculate this probability directly, because of the complexity of the integral operations in Eq. (10). Therefore, we use a simulation method to compute Eq. (10) and write the simulated log likelihood function as follows (where $v_{i,m}$ and $\eta_{i,m}$ are randomly drawn from a standard normal distribution in the m th simulation, M is the total number of simulation runs, $0 < j < J$, and n is the total number of online product reviews):

$$\text{Log likely}(\beta, \gamma, \theta, \delta) = \sum_{i=1}^n \log \frac{1}{M} \sum_{m=1}^M \left\langle \Phi \left[\frac{\mu_{i,j} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v_{i,m})}} \right] - \Phi \left[\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,i}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v_{i,m})}} \right] \right\rangle \quad (11)$$

5. Development of marginal effect-based Kano model (MEKM)

After the parameter estimation for the MOCM model, we need to interpret the resultant MOCM model to derive useful product design or improvement guidelines. In this section, we first describe the notion of the marginal effects of product features in the MOCM model. Subsequently, we propose a marginal effect-based Kano model (MEKM) to categorize customer requirements on the basis of the consumer preferences estimated by the MOCM model.

5.1. Marginal effects in the MOCM model

Model interpretation is not straightforward for ordered choice models [21,22]. Because there is no natural conditional mean function in the MOCM model, the direct interpretation of the coefficients in β is ambiguous. To assign behavioral meaning to these estimated coefficients, we use the change of $\text{Prob}[y_k = j | \mathbf{x}_{k,\cdot}, B_k, \mathbf{w}_i, \mathbf{z}_i]$, which is caused by the change of the explanatory variable $x_{k,\cdot}$ ($x_{k,\cdot} \in x_{k,\cdot}$), to compute the marginal effect:

$$\frac{\partial \text{Prob}[y_k = j | \mathbf{x}_{k,\cdot}, B_k, \mathbf{w}_i, \mathbf{z}_i]}{\partial x_{k,\cdot}} = \left[f \left(\frac{\mu_{i,j} - \beta' \cdot (\mathbf{x}_{k,\cdot}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right) - f \left(\frac{\mu_{i,j-1} - \beta' \cdot (\mathbf{x}_{k,\cdot}, B_k)}{\sqrt{\exp(\gamma' \mathbf{w}_i + v)}} \right) \right] \beta \quad (12)$$

If $x_{k,\cdot}$ is a dummy variable, its marginal effect can be computed by

$$\begin{aligned} \text{Prob}[y_k = j | \mathbf{x}_{-k,\cdot}, B_k, \mathbf{w}_i, \mathbf{z}_i; x_{k,\cdot} = 1] - \text{Prob}[y_k \\ = j | \mathbf{x}_{-k,\cdot}, B_k, \mathbf{w}_i, \mathbf{z}_i; x_{k,\cdot} = 0] \end{aligned} \quad (13)$$

where $\mathbf{x}_{-k,\cdot}$ refers to a vector containing all the regressors in the model except $x_{k,\cdot}$.

According to the data coding scheme described in Section 3.3, for product feature l , $x_{k,\cdot}$ refers to $x_{k,l,\text{pro}}$ or $x_{k,l,\text{con}}$. Both are dummy variables that indicate whether product feature l is a pro or a con in review k . The marginal effect calculated in Eq. (13) is at the individual level. After analyzing all the collected reviews and deriving the average marginal effect across all reviewers in our dataset, we obtain the aggregate marginal effect (or marginal effect for short) of $x_{l,\text{pro}}$ or $x_{l,\text{con}}$ for different rating scores. The marginal effect of $x_{l,\text{pro}}$ for rating score j , denoted $M_{l,\text{pro},j}$, means that if we improve feature l so that consumers will consider this feature a pro, the probability to give rating score j to the product will be changed by $|M_{l,\text{pro},j}|$ (the change direction depends on the sign of $M_{l,\text{pro},j}$). We interpret the meaning of $M_{l,\text{con},j}$ (the marginal effect of $x_{l,\text{con}}$ for rating score j) similarly: if we do not improve feature l so that consumers will become unsatisfied (i.e., considering this feature a con), the probability to give rating score j to the product will be changed by $|M_{l,\text{con},j}|$. Because we have a total of $J + 1$ rating levels (from 0 to J) in our model, we have $J + 1$ marginal effect pairs for each product feature extracted from online product reviews. These marginal effect pairs reflect consumers' satisfaction, at the product feature level.

5.2. Our proposed marginal effect-based Kano model (MEKM)

To categorize customer requirements, we propose the MEKM model, based on the marginal effect pairs of product features derived from the MOCM model. The marginal effect pairs (i.e., $M_{l,\text{pro},j}$ and $M_{l,\text{con},j}$ for product feature l) are defined for each rating level j (from 0 to J), so they create too many marginal effect pairs. In response, we first combine the rating levels into fewer rating groups and focus on two extremes only (i.e., low and high rating groups). We then calculate the average marginal effects for the two groups. When reviewers consider a product feature a con, it typically increases the chances that they assign low rating scores to the product. On the contrary, when reviewers become positive about a product feature (i.e., the product feature becomes a pro), it generally increases the chances that reviewers give high rating scores to that product. As a result, we concentrate on $M_{l,\text{con},\text{low}}$ and $M_{l,\text{pro},\text{high}}$ only (i.e., exclude $M_{l,\text{con},\text{high}}$ and $M_{l,\text{pro},\text{low}}$) when developing our MEKM model.

For example, in a five-star rating scale, we can combine the 1 and 2 stars to constitute the low rating group, assign 3 to represent the medium group, and use 4 and 5 stars to denote the high group. Assume that $M_{l,\text{pro},1} = 0.1$, $M_{l,\text{pro},2} = 0.2$, $M_{l,\text{pro},3} = 0.2$, $M_{l,\text{pro},4} = 0.1$, $M_{l,\text{pro},5} = 0.4$, $M_{l,\text{con},1} = 0.6$, $M_{l,\text{con},2} = 0.1$, $M_{l,\text{con},3} = 0.1$, $M_{l,\text{con},4} = 0.1$, and $M_{l,\text{con},5} = 0.1$. The average marginal effect of feature l for the low rating group when feature l is considered unsatisfactory (con) is $M_{l,\text{con},\text{low}} = 0.35$ (i.e., average of $M_{l,\text{con},1}$ and $M_{l,\text{con},2}$). That is, if we do not improve feature l and let consumers consider this feature unsatisfactory, the probability that consumers will give low rating scores (i.e., 1 and 2) to the product will increase by 0.35. The average marginal effect of feature l for the high rating group when feature l is considered satisfactory (pro) instead is $M_{l,\text{pro},\text{high}} = 0.25$ (i.e., average of $M_{l,\text{pro},4}$ and $M_{l,\text{pro},5}$). This average marginal effect indicates that if we improve feature l so that consumers are satisfied with this feature, the probability that consumers will give high rating scores (i.e., 4 and 5) to the product will increase by 0.25.

Following the general framework of the Kano model, we use $M_{l,\text{con},\text{low}}$ and $M_{l,\text{pro},\text{high}}$ to form a two-dimensional space in our proposed MEKM model, in which product features are divided into six categories according to the signs and values of the two marginal effects for each focal feature (the detailed classification conditions are in Table 1):

- 1. Must-be feature:** If both $M_{l,\text{pro},\text{high}}$ and $M_{l,\text{con},\text{low}}$ are greater than 0, but $M_{l,\text{pro},\text{high}} \ll$ (much less than) $M_{l,\text{con},\text{low}}$, then feature l is a must-be feature. That is, $M_{l,\text{con},\text{low}} > 0$ indicates that if we do not improve feature l so that consumers are negative about this feature (i.e., considering this feature a con), the probability that consumers will give low rating scores increases. On the other hand, if we improve feature l so that consumers are positive about it, the probability to assign high rating scores increases (i.e., $M_{l,\text{pro},\text{high}} > 0$); however, this increase is much less than the increase in the probability of giving low rating scores when consumers consider feature l unsatisfactory (i.e., $M_{l,\text{pro},\text{high}} \ll M_{l,\text{con},\text{low}}$). In this case, it is better to fulfill the customer requirements on product feature l .
- 2. Performance feature:** If both $M_{l,\text{pro},\text{high}}$ and $M_{l,\text{con},\text{low}}$ are greater than 0 and the values of these two variables are similar, feature l is a performance feature. In other words, the fulfillment of customer requirements on this feature will result in customer satisfaction (i.e., $M_{l,\text{pro},\text{high}} > 0$), whereas the lack of fulfillment of will prompt customer dissatisfaction, with a similar magnitude (i.e., $M_{l,\text{con},\text{low}} > 0$ and $M_{l,\text{pro},\text{high}} \approx M_{l,\text{con},\text{low}}$).
- 3. Excitement feature:** If both $M_{l,\text{pro},\text{high}}$ and $M_{l,\text{con},\text{low}}$ are greater than 0, but $M_{l,\text{pro},\text{high}} \gg M_{l,\text{con},\text{low}}$, feature l is an excitement feature. That is, the fulfillment of customer requirements on feature l will result in customer satisfaction (i.e., $M_{l,\text{pro},\text{high}} > 0$), but the lack of fulfillment of customer requirements on feature l causes dissatisfaction (i.e., $M_{l,\text{con},\text{low}} > 0$) with a lesser degree compared to $M_{l,\text{pro},\text{high}}$ (i.e., $M_{l,\text{pro},\text{high}} \gg M_{l,\text{con},\text{low}}$).

Table 1
Classification conditions in the MEKM model.

Condition	Feature category
$M_{l,\text{pro},\text{high}} > 0$, $M_{l,\text{con},\text{low}} > 0$ and $M_{l,\text{pro},\text{high}} \ll M_{l,\text{con},\text{low}}$	Must-be feature
$M_{l,\text{pro},\text{high}} > 0$, $M_{l,\text{con},\text{low}} > 0$ and $M_{l,\text{pro},\text{high}} \approx M_{l,\text{con},\text{low}}$	Performance feature
$M_{l,\text{pro},\text{high}} > 0$, $M_{l,\text{con},\text{low}} > 0$ and $M_{l,\text{pro},\text{high}} \gg M_{l,\text{con},\text{low}}$	Excitement feature
$M_{l,\text{pro},\text{high}} < 0$, $M_{l,\text{con},\text{low}} > 0$	Innovation-needed feature
$M_{l,\text{pro},\text{high}} < 0$, $M_{l,\text{con},\text{low}} < 0$	Reverse feature
$M_{l,\text{pro},\text{high}} > 0$, $M_{l,\text{con},\text{low}} < 0$	Divergent feature

4. **Innovation-needed feature:** If $M_{I,pro,high}$ is less than 0, and $M_{I,con,low}$ is greater than 0, feature I is an innovation-needed feature. That is because, in this case, whether we improve feature I or not, consumers will increase their probability to assign low rating scores (i.e., $M_{I,con,low} > 0$) and decrease their probability to offer high rating scores (i.e., $M_{I,pro,high} < 0$). In other words, consumers are not satisfied with the current design of feature I , so a revolutionary change in this feature I may be welcomed.
5. **Reverse feature:** If both $M_{I,pro,high}$ and $M_{I,con,low}$ are less than 0, then feature I is a reverse feature. This condition denotes that the fulfillment of feature I will result in customer dissatisfaction (i.e., $M_{I,pro,high} < 0$), whereas leaving feature I unfulfilled will invoke customer satisfaction (i.e., $M_{I,con,low} < 0$).
6. **Divergent feature:** If $M_{I,pro,high}$ is greater than 0 but $M_{I,con,low}$ is less than 0, feature I is a divergent feature. In this case, the fulfillment of feature I will result in customer satisfaction (i.e., $M_{I,pro,high} > 0$), but so does its lack of fulfillment (i.e., $M_{I,con,low} < 0$). That is, customers' opinions about feature I are conflicting.

By using a two-dimensional graph in which the horizontal axis represents $M_{I,pro,high}$ and the vertical axis denotes $M_{I,con,low}$, we can map these six product feature categories to six areas (Fig. 3). The performance feature lies around the 45 degree dashed line (i.e., $M_{I,pro,high} \approx M_{I,con,low}$) in the first quadrant (i.e., where $M_{I,pro,high} > 0$ and $M_{I,con,low} > 0$).

6. Empirical study

In this section, we conduct an empirical study to evaluate the effectiveness of the proposed econometric preference measurement model (i.e., MOCM) and demonstrate the utility of our proposed MEKM model, using online product reviews collected from a popular product review website, Epinions.com. In the following, we describe our data collection and then some important evaluation results.

6.1. Data collection and preparation

We collected all product reviews about mobile phones posted on Epinions.com before November 2103, together with reviewer profiles. Each online product review included the title of the review, pro/con phrases, the detailed free-text review, the date

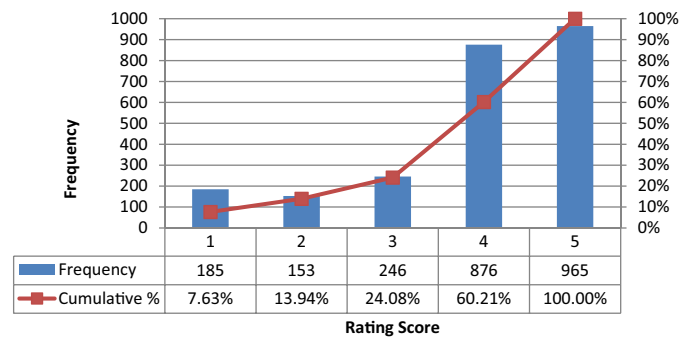


Fig. 4. Distribution of rating scores in our dataset.

Table 2
Statistics of our dataset.

	Number of reviews by a reviewer	Number of members trusting a reviewer	Number of members trusted by a reviewer
Min	1	0	0
Max	18	450	446
Average	1.084	8.183	7.832
Standard deviation	0.5524	35.9755	34.5049

that the review was posted, and the overall product rating (on a five-star scale). The reviewer profile consisted of the list of reviews this reviewer had written and his or her trust network (i.e., who trusts the focal reviewer and who is trusted by this reviewer). Because mobile phones typically invoke plentiful discussions online, we considered any phones with few online reviews are not well discussed, such that the opinions expressed in this limited set of reviews likely are not representative. Accordingly, we discarded brands that prompted no more than 55 reviews.³ Finally, we obtained a dataset with 2425 online product reviews, spanning 9 popular brands and 186 models in the mobile phone market.

Fig. 4 illustrates the distribution of different rating scores in our dataset. This positively skewed (J-shaped) distribution is consistent with the pattern commonly found on the online rating platforms [14,41]. Table 2 shows some statistics of our dataset, including the number of reviews written by a reviewer, the number of members trusting a reviewer, and the number of members trusted by a reviewer.

We extracted 25 features from the pro/con phrases across all product reviews in our dataset by using the product feature extraction procedure (denoted as the PFE procedure) described in Section 3.1. To assess the quality of these extracted product features, we conducted a validation experiment. In this experiment, we randomly drew two non-overlapped samples from our dataset, and each sample contained 100 product reviews. Subsequently, we recruited two graduate students from a business school, both with at least six years' experience with mobile phone usage. Each coder was asked to examine one sample (i.e., 100 reviews) by assigning each pro and con mentioned in these reviews into the 25 product features identified by the PFE procedure. On the basis of these two samples, we compared the product features extracted by the PFE procedure with those manually assigned by the two coders. As Table 3

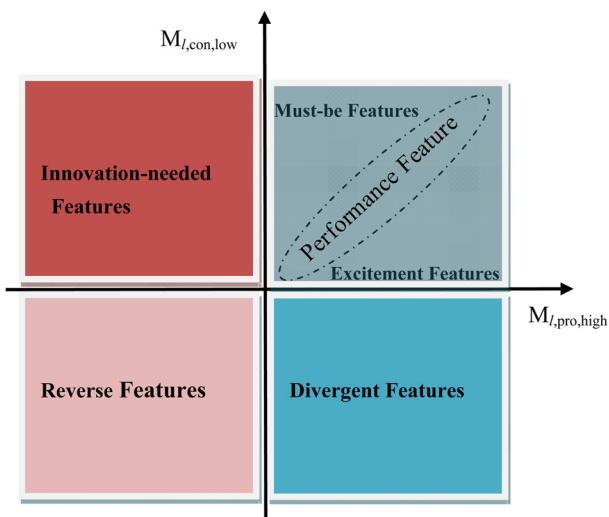


Fig. 3. Product feature classification in the proposed MEKM model.

³ Our original dataset covers 11 mobile brands. For each brand, we counted the number of reviews discussing the products pertaining to that brand. After sorting the 11 brands according to their number of reviews, we found that the three brands with least number of reviews had 105, 55, and 46 reviews, respectively. Considering the gap between 105 and 55, we decided to choose 55 as the threshold to discard brands.

Table 3
Result of the validation experiment.

Feature	Description	Frequency (%)	Feature	Description	Frequency (%)
Apps	Apps and games	8.25 [†] (9.00 [‡])	Multimedia	Multimedia/music and video	3.25 (4.00)
Battery	Durability, weight, and charging efficiency of battery	55.50 (55.50)	OS	Operation system and reliability	2.25 (3.75)
Camera	Camera and photo quality	17.75 (17.75)	Price	Price	8.75 (9.00)
Design	Design, quality, and material of outer casing	9.25 (11.25)	Reliability	Mobile phone's reliability	24.75 (18.00)
Display	Display brightness and quality	4.75 (6.25)	Screen	Screen resolution and type	19.00 (19.25)
Easy to use	Easy to use	13.75 (16.50)	Appearance	Mobile phone's size and shape	9.50 (11.50)
Navigation	GPS and navigation support	2.50 (2.75)	Software	Built-in software	5.00 (5.50)
Interface	Interface/inner equipment	16.25 (15.70)	Speed	Mobile phone's response speed	10.25 (6.50)
Web browsing	Internet/web browsing	6.50 (4.75)	Voice	Voice quality and volume	8.75 (6.75)
Configuration	Memory/processor	6.25 (8.75)	Weight	Weight	5.25 (4.00)
Text support	Message/email/texting support	5.50 (5.75)	Wifi	Wifi support	1.75 (1.75)
Accessory	Mobile accessory	5.00 (3.25)	New functions	Video telephone, speech to text and other new functions	18.75 (19.50)
Service	After sales services	3.25 (7.75)			

[†] Percentage of pros and cons in the two samples (i.e., 200 reviews) classified into this product feature by the product feature extraction procedure (Section 3.1).

[‡] Percentage of pros and cons in the two samples classified into this product feature by the coders.

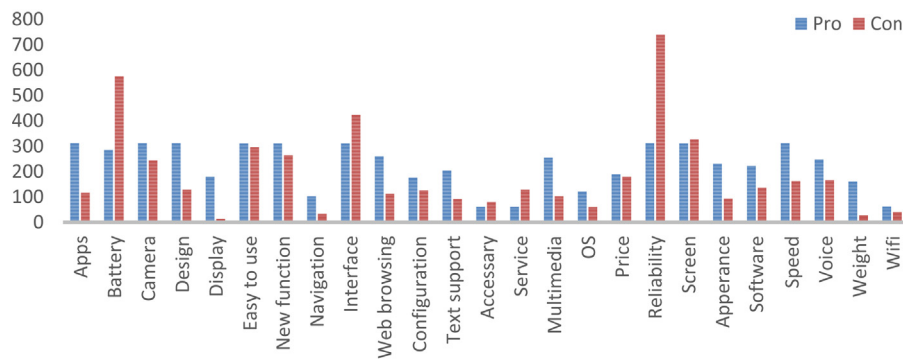


Fig. 5. Distribution of features discussed in the pro and con fields of the reviews.

illustrates, their relative frequencies were comparable across all product features. The Pearson correlation coefficient was 0.98, suggesting the effectiveness of our PFE procedure. We also calculated the Cohen's kappa coefficient for each product feature to measure the agreement between the two methods (i.e., PFE and manual assignment).⁴ The average Cohen's kappa coefficient across the 25 product features was 0.8129, with a maximum of 1 ("battery") and a minimum of 0.6468 ("reliability and quality"). The Cohen's kappa coefficient indicated good agreement between the two methods, which also suggested the satisfactory effectiveness of our PFE procedure.

Fig. 5 shows the distributions of the 25 features in our dataset. The "Apps," "design," and "weight" features were more frequently considered pros rather than cons, whereas reviewers' sentiments on "easy to use," "price," and "screen" features were quite balanced. In contrast, reviewers were more negative on the "battery," "interface," and "reliability" features.

Subsequently, we used the trust relations between reviewers in our dataset to construct a trust network (Fig. 6). We then employed the normalized PageRank value [35] of each reviewer in this network to represent his or her position or reputation in the community. When calculating PageRank values, the damping factor was set as 0.85. We also normalized the total number of

reviews written by each reviewer, using $(n_i - \min(\cdot)) / (\max(\cdot) - \min(\cdot))$, where n_i denotes the number of reviews written by reviewer i , and $\min(\cdot)$ and $\max(\cdot)$ refer to the minimum and maximum number of reviews written among all reviewers in the dataset, respectively.

6.2. Model estimation and comparison

In this section, we use the collected dataset to estimate our econometric preference measurement model (i.e., MOCM). We also

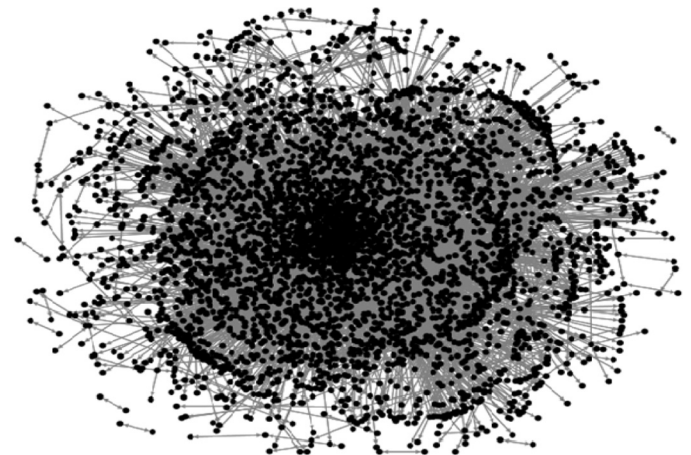


Fig. 6. Trust network of reviewers.

⁴ The Cohen's kappa coefficient, which takes a value between -1 and 1 , is widely used to measure the agreement between two raters, each of whom classifies N items into C mutually exclusive categories. Values less than 0 indicate no agreement, 0 – 0.20 slight, 0.21 – 0.40 fair, 0.41 – 0.60 moderate, 0.61 – 0.80 substantial, and 0.81 – 1 almost perfect agreement.

estimate aggregate consumer preferences using three prevalent models as the benchmarks for our proposed model: Poisson Regression Model (PRM), Negative Binomial Regression Model (NBRM), and Basic Ordered Choice Model (BOCM).

For both PRM and NBRM, we follow Decker and Trusov [13] and employ the 25 product features extracted from online product reviews, in conjunction with reviewers' sentiment orientations (pros or cons) toward the features, as the independent variables (i.e., a total of 50 independent variables) and the nine brand names as control variables. Furthermore, we assume that the rating score (i.e., the dependent variable) satisfies a Poisson distribution for the PRM model or a negative binominal distribution for the NBRM model. Accordingly, PRM and NBRM can be described by Eqs. (14) and (15), respectively:

$$\text{Prob}[y_k = j | \mathbf{x}_{k,\cdot}, B_k] = \frac{\lambda^j}{j!} \exp(-\lambda), \quad \text{where } \lambda = \exp(\alpha + \beta' \cdot (\mathbf{x}_{k,\cdot,\text{pro}}, \mathbf{x}_{k,\cdot,\text{con}}, B_k)) \quad (14)$$

$$\text{Prob}[y_k = j | \mathbf{x}_{k,\cdot}, B_k] = \binom{j + \tau - 1}{j} \frac{\lambda^j \tau^\tau}{(\lambda + \tau)^{j+\tau}}, \quad \text{where } \tau > 0 \text{ and } \lambda = \exp(\alpha + \beta' \cdot (\mathbf{x}_{k,\cdot,\text{pro}}, \mathbf{x}_{k,\cdot,\text{con}}, B_k)) \quad (15)$$

In Eq. (15), τ can be taken as the Gamma parameter in the distribution. All the parameters in Eqs. (14) and (15) are estimated using the maximum likelihood methods.

For the BOCM model (see Eqs. (1) and (2)), though it uses the same independent and control variables as PRM and NBRM do, the dependent variable (the rating score) is modeled by a more reasonable ordinal model. Finally, in our proposed MOCM model, we extend the BOCM model to account for reviewers' heterogeneity. The brand name and reviewer's sentiment orientations toward each product feature are treated as explanatory variables (independent variables; Eq. (7)). To account for reviewer heterogeneity, we also consider the number of reviews each reviewer has written (n_i) after normalization and the normalized PageRank value of each reviewer (tr_i). In our empirical study, the former is treated as an explanatory variable for Eq. (4) and the latter is an explanatory variable in Eq. (6).⁵ The parameters in the MOCM model are estimated by the maximum simulated likelihood method.

What should be noted here is that, we do not incorporate reviewer profile (i.e., tr_i and n_i) into the PRM and NBRM models when comparing the performance of different models. That is mainly because these two models are counting models and are usually designed to model interval scale data rather than ordinal scale data. Furthermore, unlike product features ($\mathbf{x}_{k,\cdot,\text{con}}, \mathbf{x}_{k,\cdot,\text{pro}}$) and product brand (B_k), individual reviewer profiles do not affect rating scores directly. Instead, they generally influence rating scores indirectly through reviewers' rating behavior. As a result, it is not reasonable to directly incorporate this information into PRM and NBRM. Incorporating reviewer profiles indirectly into PRM and NBRM and comparing their performance with our proposed MOCM model will be our future research work.

Table 4 illustrates the estimation results of the three benchmark models and our proposed MOCM model. According

to Table 4, most product features with significant effects have the expected signs (i.e., pros with a positive sign and cons with a negative sign). The signs for the control variables (brands) are mostly consistent across the four models. Table 5 summarizes the comparison results for these four models. In this study, we use the log likelihood and Akaike Information Criterion (AIC) [2] as the model evaluation criteria. The log likelihood is a criterion that measures the goodness of fit of a given statistical (or econometric) model, whereas AIC is a relative estimate of the information loss when the model is used to represent the "true" model that generates a given set of data. The information loss estimated by AIC takes into account the tradeoff between the goodness of fit and the complexity of the model. Given a set of candidate models, the model that minimizes the information loss (i.e., has the lowest AIC value) is considered the best model. The log likelihood measure may be biased toward a more complex model, but AIC does not because it also considers the model complexity. As Table 5 illustrates, among the four models examined, our proposed MOCM model attains the highest value in the log likelihood measure, but its AIC value is the lowest. Thus, in terms of log likelihood and AIC, our proposed MOCM model outperforms the extant three preference measurement models (i.e., PRM, NBRM, and BOCM).

However, a model with high explanatory power may be subject to potential predictive falsification [13,51]. To test for this concern, we conduct 15 random and independent cross-validation experiments by using the holdout method to compare the hit rate of different models. Specifically, in each experiment, we first randomly divide the dataset into a calibration set (containing 90% of the dataset, i.e., 2183 observations) and a prediction set (10% of the dataset, i.e., 242 observations). Subsequently, we use the calibration set to train a model and the prediction set to test its prediction effectiveness (measured by hit rate). We repeat this process 15 times and the overall performance of the model is the average of the 15 trials. Fig. 7 shows the results of the 15 experiments for all models. As Fig. 6 illustrates, our proposed MOCM model consistently outperforms the three benchmark models over the 15 trials, and the PRM model again represents the least effective model (measured by hit rate). The t -test ($\alpha = 0.05$ and $\alpha = 0.1$) between the proposed MOCM model and each benchmark model also supports our conclusion.

We further perform a sensitivity analysis on the granularity of the rating scale (i.e., dependent variable). In the previous evaluations, the rating scale consists of 5 levels; with this experiment, we transform the scale into three levels (i.e., low, medium, and high) and examine whether our proposed MOCM model still performs better than the three benchmark models. The results show that the hit rate of each model improves. In addition, our proposed MOCM model still outperforms the other three models. Specifically, the hit rate of the MOCM model is 61.30%, which is greater than that for PRM (49.56%), NBRM (55.86%), or BOCM (56.25%). This sensitivity analysis accordingly demonstrates the robustness of our proposed econometric preference measurement model (i.e., MOCM).

6.3. Categorization of customer requirements using MEKM

Using the MOCM model estimated in the previous section, we calculate, for each product feature, average marginal effect pairs for the high and low rating groups, then analyze the customer requirements using our proposed MEKM model. Fig. 8 illustrates the results of the MEKM model, which maps the 25 product features onto a two-dimensional graph. Six product features are considered performance features: "navigation," "multimedia," "screen," "Web browsing," "camera," and "voice." "Reliability," "text support," and "easy to use" are labeled as must-be features; "price," "weight," "battery," "OS (operating system)," "configuration," "Wifi," "speed,"

⁵ We conducted a series of experiments to determine an appropriate model for explaining reviewer heterogeneity in the MOCM model. Specifically, we examined four models: (1) both n_i and tr_i are explanatory variables in Eq. (4); (2) both n_i and tr_i are explanatory variables in Eq. (6); (3) n_i is an explanatory variable in Eq. (4) and tr_i is an explanatory variable in Eq. (6); and (4) n_i is an explanatory variable in Eq. (6) and tr_i as an explanatory variable in Eq. (4). The experimental results suggest that the third model achieved the best model fitness, so we adopted it and treated n_i and tr_i as explanatory variables in Eqs. (4) and (6), respectively.

Table 4

Parameter estimates of different empirical models.

Variable	PRM		NBRM		BOCM		MOCM	
	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error	Estimate	Std. error
Intercept	1.346***	0.045	1.348***	0.046	1.533***	0.153	6.045***	0.886
Pro_Apps	0.042***	0.014	0.045***	0.016	0.170**	0.074	0.698*	0.407
Pro_Battery	0.043***	0.016	0.049***	0.017	0.177**	0.074	0.947**	0.396
Pro_Camera	0.034**	0.015	0.038**	0.019	0.146*	0.080	0.863*	0.464
Pro_Design	−0.021	0.016	−0.032	0.516	−0.108	0.072	−0.439	0.366
Pro_Display	−0.02	0.023	−0.023	0.123	−0.067	0.087	−0.265	0.391
Pro_Easy to use	0.076***	0.014	0.078***	0.019	0.231***	0.074	1.557***	0.488
Pro_New functions	0.075***	0.016	0.076***	0.021	0.297***	0.071	1.295***	0.387
Pro_Navigation	0.057**	0.023	0.061**	0.028	0.236**	0.117	1.094*	0.628
Pro_Interface	0.056**	0.019	0.060**	0.021	0.194**	0.074	0.882**	0.354
Pro_Web browsing	0.030*	0.017	0.032*	0.018	0.086	0.079	0.345	0.409
Pro_Configuration	0.0489***	0.019	0.053***	0.021	0.193**	0.093	1.289**	0.587
Pro_Text support	−0.009	0.022	−0.012	0.124	−0.062	0.085	−0.114	0.413
Pro_Accessary	−0.004	0.037	−0.006	0.069	−0.03	0.146	−0.191	0.726
Pro_Service	0.050	0.034	0.050	0.041	0.164	0.149	1.059	0.834
Pro_Multimedia	0.035**	0.017	0.035**	0.018	0.126	0.078	0.676*	0.411
Pro_OS	0.067**	0.019	0.071**	0.021	0.331***	0.113	1.340**	0.668
Pro_Price	−0.013	0.022	−0.013	0.029	−0.097	0.086	−0.451	0.421
Pro_Reliability	0.061***	0.016	0.068***	0.021	0.198***	0.073	0.930**	0.386
Pro_Screen	0.040**	0.017	0.040**	0.018	0.174**	0.077	0.861*	0.410
Pro_Appearance	0.036*	0.020	0.037*	0.021	0.168**	0.080	0.702*	0.399
Pro_Software	−0.001	0.019	−0.001	0.025	−0.016	0.084	0.142	0.424
Pro_Speed	0.106***	0.014	0.122***	0.019	0.491***	0.078	2.952***	0.641
Pro_Voice	0.015	0.018	0.015	0.018	0.065	0.077	0.431	0.376
Pro_Weight	0.043**	0.021	0.048**	0.023	0.189**	0.094	0.711	0.481
Pro_Wifi	0.087***	0.031	0.091***	0.034	0.358**	0.149	2.241**	0.945
Con_Apps	0.053**	0.020	0.079**	0.028	0.195*	0.111	1.142*	0.651
Con_battery	−0.035*	0.013	−0.037*	0.015	−0.210***	0.057	−0.889***	0.298
Con_Camera	0.022	0.017	0.022	0.048	0.022	0.077	0.454	0.417
Con_Design	−0.045	0.029	−0.045	0.069	−0.232**	0.103	−0.514	0.557
Con_Display	−0.031	0.060	−0.031	0.08	−0.311	0.309	−1.384	1.554
Con_Easy to use	−0.054***	0.021	−0.079***	0.029	−0.156**	0.070	−0.938**	0.327
Con_New functions	−0.058***	0.022	−0.083***	0.031	−0.240***	0.079	−1.360***	0.389
Con_Navigation	−0.087	0.066	−0.087	0.166	−0.278	0.207	−1.071	0.880
Con_Interface	−0.025	0.016	−0.025	0.086	−0.107*	0.061	−0.528*	0.285
Con_Web browsing	0.009	0.026	0.009	0.126	−0.029	0.113	−0.175	0.554
Con_Configuration	−0.003	0.023	−0.003	0.068	−0.070	0.104	−0.181	0.533
Con_Text support	−0.106***	0.033	−0.146***	0.038	−0.359***	0.117	−1.78***	0.577
Con_Accessary	0.021	0.028	0.023	0.126	0.111	0.134	0.547	0.657
Con_Service	0.022	0.025	0.024	0.167	0.110	0.106	0.308	0.488
Con_Multimedia	−0.007	0.026	−0.087	0.226	−0.031	0.117	−0.518	0.532
Con_OS	−0.033	0.040	−0.035	0.15	−0.149	0.152	−0.661	0.713
Con_Price	−0.004	0.020	−0.004	0.19	−0.019	0.092	−0.028	0.447
Con_Reliability	−0.091***	0.013	−0.121***	0.019	−0.365***	0.052	−1.799***	0.332
Con_Screen	−0.050***	0.017	−0.075***	0.021	−0.227***	0.068	−1.065***	0.357
Con_Appearance	0.049**	0.023	0.050**	0.024	0.186	0.132	0.986	0.860
Con_Software	−0.046	0.029	−0.043	0.031	−0.179*	0.100	−0.759*	0.452
Con_Speed	−0.121***	0.027	−0.141***	0.03	−0.471***	0.092	−1.778***	0.497
Con_Voice	−0.026	0.024	−0.026	0.037	−0.154*	0.091	−0.767*	0.437
Con_Weight	0.009	0.046	0.007	0.049	0.019	0.225	0.311	1.494
Con_Wifi	0.022	0.035	0.025	0.612	0.034	0.179	−0.108	0.896
Brand 1	0.074	0.046	0.094	0.546	0.435***	0.162	2.213***	0.772
Brand 2	−0.035	0.047	−0.043	0.147	−0.080	0.157	−0.285	0.667
Brand 3	−0.009	0.050	−0.009	0.15	0.002	0.182	−0.184	0.812
Brand 4	−0.111**	0.052	−0.113**	0.054	−0.358**	0.175	−1.490*	0.777
Brand 5	−0.213***	0.053	−0.263***	0.069	−0.613***	0.171	−2.506***	0.778
Brand 6	0.053	0.046	0.053	0.176	0.231	0.164	1.009	0.726
Brand 7	−0.018	0.057	−0.018	0.527	0.007	0.196	−0.070	0.827
Brand 8	0.064	0.045	0.074	0.845	0.241	0.160	1.308*	0.737
Brand 9	–	–	–	–	–	–	–	–
NORRW	–	–	–	–	–	–	−0.014*	0.024
NORPR	–	–	–	–	–	–	0.040***	0.014
Log-likelihood	−4313.61		−4128.53		−2986.90		−2961.10	
Akaike Information Criterion	8745.20		8375.06		6097.80		6066.20	

* $p < 0.05$.** $p < 0.01$.*** $p < 0.001$.

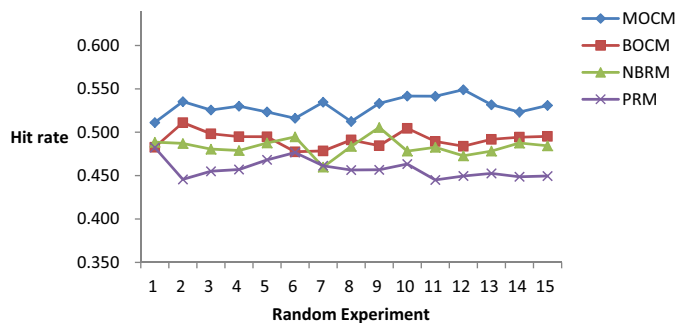
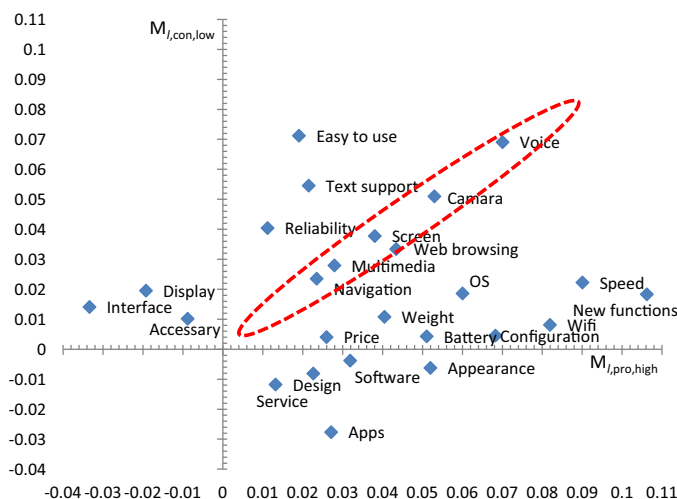
and “new functions” are classified as excitement features. On the other hand, the divergent features include “service,” “design,” “Apps,” “software,” and “appearance,” whereas innovation-needed features encompass “accessory,” “display,” and “interface.” No product feature is classified as a reverse feature.

The managerial implications from the MEKM model can be derived as follows. Must-be features (“reliability,” “text support,” and “easy to use”) represent customers’ basic requirements for a mobile phone. Thus, all mobile phone companies must ensure that their mobile phones fulfill these basic requirements. Furthermore,

Table 5

Comparison of different models.

		PRM	NBRM	BOCM	MOCM
Property	Heterogeneity	No	Yes	No	Yes
	Heteroscedasticity	–	–	No	Yes
	Different thresholds	–	–	No	Yes
Model fitness	Log likelihood	–4313.61	–4128.53	–2986.90	–2961.10
	Akaike Information Criterion	8745.20	8375.06	6097.80	6066.20

**Fig. 7.** Hit rates of different models.**Fig. 8.** Classification of features of mobile phones using the MEKM model.

the MEKM model can help mobile phone companies to develop their product design or improvement plans according to their competitive strategies.⁶ For example, mobile phone companies with a cost leadership strategy typically target a broad market rather than a narrow, niche market and make their efforts to reduce product costs to become low-cost producers in the industry. Because customer satisfaction is proportional to the level of fulfillment of performance features (i.e., “navigation,” “multimedia,” “screen,” “Web browsing,” “camera,” and “voice”), most mobile phones compete on the performance features. Thus, the market for mobile phones that focus on achieving performance features generally is a relatively broad segment. Mobile phone companies that pursue a cost leadership strategy therefore should compete on this market and develop cost-effective methods to fulfill these performance features.

Mobile phone companies with a differentiation strategy instead should focus on excitement features (“price,” “weight,” “battery,” “OS (operating system),” “configuration,” “WIFI,” “speed,” and

“new functions”). Customers often can tolerate the current design of an excitement feature (even if their requirements are not completely fulfilled), but will be delighted when this feature is improved. Thus, excitement features offer a wide range of possibilities for differentiation. In addition, companies with a differentiation strategy can also concentrate on the divergent features (“service,” “design,” “Apps,” “software,” and “appearance”) and find ways to differentiate their products from competitors’ ones. Because consumer preferences for divergent features differ or even conflict, firms can attempt to group consumer preferences with respect to the divergent features and design different types of mobile phones to match those varying preferences and seek new market segments.

Finally, if mobile phone companies adopt an innovation strategy, they should pay more attention to innovation-needed features (“accessory,” “display,” and “interface”), because customers are not satisfied with their current designs. Thus, these product features call for innovative designs, and companies should target some or all of the innovation-needed features by designing mobile phones with revolutionary changes on the selected features.

7. Conclusions and future research directions

Opinion sharing platforms provide open, convenient communication channels for sharing and gathering consumers’ experiences with and preferences for various products. Meanwhile, these online product reviews also represent unique and valuable information sources for firms to understand the preferences of their customers. In this study, we focus on how to measure aggregate consumer preferences from online product reviews and then categorize customer requirements for supporting product design/improvement decisions. Accordingly, we first establish a framework for semi-automatically extracting product features and reviewers’ sentiment orientations from online product reviews. We then propose an econometric preference measurement model, i.e., the modified ordered choice model (MOCM), to extract aggregate consumer preferences. Furthermore, to categorize customer requirements and support product design/improvement at the product feature level, we propose a marginal effect-based Kano model (MEKM) by extending the Kano model.

Our research contributions are three-fold. First, this study contributes to online product review research in information systems (IS) field. Most of IS literature on product reviews investigate how online product reviews affect product sales [9,17,34,61] or examine the antecedents of the helpfulness of product reviews [33,44,60]. Our current study extends the scope of IS research on online product reviews and propose effective and viable methods for converting a vast amount of data (i.e., online product reviews) into useful business intelligence (i.e., preference measurement).

Second, this study also contributes to the preference measurement literature by offering a more effective econometric model with a strong theoretical foundation to estimate aggregate consumer preferences from online product reviews. Traditionally, aggregate consumer preferences are estimated by means of

⁶ According to Porter and Millar [50], firms’ competitive strategies can be divided into four categories: cost leadership, differentiation, innovation, and growth.

conjoint analysis using preference data collected from surveys or experiments. However, such preference elicitation method is usually time consuming and costly. This study provides an econometric framework that turns online product reviews into aggregated consumer preferences. Specifically, we develop a modified ordered choice model (MOCM), which considers the heteroscedasticity of reviewers' rating variance and allows reviewers to assign rating scores according to their own thresholds. Using a dataset collected from a popular product review website, Epinions.com, our empirical study shows that the proposed MOCM model outperforms the three benchmark models (i.e., PRM, NBRM, and BOCM).

Third, our study also contributes to the product design literature by proposing a marginal effect-based Kano model (MEKM) to categorize and prioritize customer requirements at the product feature level. The proposed MEKM model seamlessly integrates our proposed econometric preference measurement model and the traditional Kano model to analyze and visually portray the impact of product feature improvement on consumer satisfaction. Using our dataset, we empirically demonstrate the utility of the MEKM model.

Our research also has several practical implications. First, this study offers useful models (i.e., MOCM and MEKM) that managers and practitioners can use to extract and measure aggregate consumer preferences and categorize customer requirements from online product reviews. For example, product designers can employ our proposed MEKM model to comprehend customer requirements by categorizing product features into different categories, including must-be, performance, excitement, innovation-needed, divergent, and reverse features. Such understanding can facilitate product designers to derive product design or improvement plans according to their competitive strategies. Second, marketing managers can also exploit the categorization of product features for improving the effectiveness of their advertising decisions. For example, for firms with a differentiation strategy, their advertising should emphasize excitement and divergent features (if they have improved these features) rather than highlighting performance features.

Our study has some limitations that warrant additional research attention. First, in this study, we assume that all product reviews available on online product review platforms are contributed by honest reviewers and, thus, their reviews are not manipulated or fake. However, spam reviews can easily be found online. How to identify such untruthful reviews is a growing research issue [29,45]. Combining our study with a good spam review detection method would constitute an important, interesting research direction. Second, our current study only focuses on semi-structured reviews. However, as we mentioned previously, our research can easily be extended to handle free-text reviews. Thus, to enhance the applicability of our proposed econometric framework, it is desirable to incorporate an existing opinion mining method or develop a new opinion mining method such that our proposed framework can also take free-text reviews as inputs. Third, a possible concern of our study stems from the representativeness of reviewers in online product review websites. It is commonly known that online reviewers may not be representative of the target population. Therefore, the appropriateness of using online product reviews to infer general consumer preferences depends on the representativeness of the reviewers and online product reviews included in the analysis. To lessen this concern, a possible direction for future research is to include reviews from multiple product review websites in order to reduce biases associated with individual websites. Alternatively, we can also collect the profiles of reviewers and adopt some statistical methods to control the possible self-selection problem in online review data. Finally, a promising future research direction is to add

a temporal dimension to the current study. It will be interesting to develop an econometric model to capture the evolution of aggregate consumer preferences and then examine how aggregate consumer preferences evolve over time and how the evolution of consumer preferences influences product design or improvement decisions.

Acknowledgments

This work was supported in part by the Ministry of Science and Technology of Taiwan under the grant NSC under the grants NSC 100-2410-H-002-021-MY3 and NSC 101-2410-H-002-041-MY3 and by the Research Fund for the Doctoral Program of Higher Education of China (20120073110029), Inter Discipline Foundation of Shanghai Jiao Tong University (No. 11JCZ02), National Natural Science Foundation of China (Nos. 71371123, 71131005), and Europe-China High Value Engineering Network (EC-HVEN: 295130).

References

- [1] A.S. Abrahams, J. Jiao, W. Fan, G.A. Wang, Z. Zhang, What's buzzing in the blizzard of buzz? Automotive component isolation in social media postings *Decis. Support Syst.* 55 (4), 2013, pp. 871–882.
- [2] H. Akaike, A new look at the statistical model identification, *IEEE Trans. Autom. Control* 19 (6), 1974, pp. 716–723.
- [3] N. Archak, A. Ghose, P.G. Ipeirotis, Deriving the pricing power of product features by mining consumer reviews, *Manage. Sci.* 57 (8), 2011, pp. 1485–1509.
- [4] N. Archak, A. Ghose, P.G. Ipeirotis, Show me the money! Deriving the pricing power of product features by mining consumer reviews in: *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2007, pp. 56–65.
- [5] C. Berger, R. Blauth, D. Boger, C. Bolster, G. Burchill, W. DuMouchel, F. Pouliot, R. Richter, A. Rubinoff, D. Shen, M. Timko, D. Walden, Kano's methods for understanding customer-defined quality, *Center Qual. Manag. J.* 2 (4), 1993, pp. 3–35.
- [6] S. Branavan, H. Chen, J. Eisenstein, R. Barzilay, Learning document-level semantic properties from free-text annotations, *J. Artif. Intell. Res.* 34 (2), 2009, pp. 569–603.
- [7] E. Cambria, B. Schuller, Y. Xia, C. Havasi, New avenues in opinion mining and sentiment analysis, *IEEE Intell. Syst.* 28 (2), 2013, pp. 15–21.
- [8] C.C. Chen, M.C. Chuang, Integrating the Kano model into a robust design approach to enhance customer satisfaction with product design, *Int. J. Prod. Econ.* 114 (2), 2008, pp. 667–681.
- [9] Y. Chen, J. Xie, Online consumer review: word-of-mouth as a new element of marketing communication mix, *Manage. Sci.* 54 (3), 2008, pp. 477–491.
- [10] Y. Chen, Q. Wang, J. Xie, Online social interactions: a natural experiment on word of mouth versus observational learning, *J. Mark. Res.* 48 (2), 2011, pp. 238–254.
- [11] J.A. Chevalier, D. Mayzlin, The effect of word of mouth on sales: online book reviews, *J. Mark. Res.* 43 (3), 2006, pp. 345–354.
- [12] W. Dai, D. Han, Y. Dai, D. Xu, Emotion recognition and affective computing on vocal social media, *Inform. Manag.* 52 (7), 2015, pp. 777–788.
- [13] R. Decker, M. Trusov, Estimating aggregate consumer preferences from online product reviews, *Int. J. Res. Mark.* 27 (4), 2010, pp. 293–307.
- [14] C. Dellarocas, R. Narayan, A statistical measure of a population's propensity to engage in post-purchase online word-of-mouth, *Stat. Sci.* 21 (2), 2006, pp. 277–285.
- [15] N. Eluru, C.R. Bhat, D.A. Hensher, A mixed generalized ordered response model for examining pedestrian and bicyclist injury severity level in traffic crashes, *Accid. Anal. Prev.* 40 (3), 2008, pp. 1033–1054.
- [16] P.S. Fader, B.G. Hardie, Modeling consumer choice among SKUs, *J. Mark. Res.* 33 (4), 1996, pp. 442–452.
- [17] C. Forman, A. Ghose, B. Wiesenfeld, Examining the relationship between reviews and sales: the role of reviewer identity disclosure in electronic markets, *Inform. Syst. Res.* 19 (3), 2008, pp. 291–313.
- [18] A. Ghose, P.G. Ipeirotis, A. Sundararajan, Opinion mining using econometrics: a case study on reputation systems, in: *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, 2007, pp. 416–423.
- [19] A. Ghose, P.G. Ipeirotis, B. Li, Designing ranking systems for hotels on travel search engines by mining user-generated and crowd sourced content, *Mark. Sci.* 31 (3), 2012, pp. 493–520.
- [20] P.E. Green, V.R. Rao, Conjoint measurement for quantifying judgmental data, *J. Mark. Res.* 8 (3), 1971, pp. 355–363.
- [21] W.H. Greene, D.A. Hensher, *Modeling Ordered Choices: A Primer*, Cambridge University Press, 2010.
- [22] W. Greene, M.N. Harris, B. Hollingsworth, T.A. Weterings, Heterogeneity in ordered choice models: a review with applications to self-assessed health, *J. Econ. Surv.* 28 (1), 2014, pp. 109–133.
- [23] B. Gu, J. Park, P. Konana, The impact of external word-of-mouth sources on retailer sales of high-involvement products, *Inform. Syst. Res.* 23 (1), 2012, pp. 182–196.
- [24] M. Halme, M. Kallio, Estimation methods for choice-based conjoint analysis of consumer preferences, *Eur. J. Oper. Res.* 214 (1), 2011, pp. 160–167.

- [25] W. He, H. Wu, G. Yan, V. Akula, J. Shen, A novel social media competitive analytics framework with sentiment benchmarks, *Inform. Manag.* 52 (7), 2015, pp. 801–812.
- [26] M. Hu, B. Liu, Mining and summarizing customer reviews, in: *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2004, pp. 168–177.
- [27] S.K. Hui, P.S. Fader, E.T. Bradlow, Path data in marketing: an integrative framework and prospectus for model building, *Mark. Sci.* 28 (2), 2009, pp. 320–335.
- [28] H. Janne, P. Timo, Sharpening logistics customer service strategy planning by applying Kano's quality element classification, *Int. J. Prod. Econ.* 56/57, 1998, pp. 253–260.
- [29] N. Jindal, B. Liu, Opinion spam and analysis, in: *Proceedings of the 2008 International Conference on Web Search and Data Mining*, 2008, pp. 219–230.
- [30] N. Kano, N. Seraku, F. Takahashi, S. Tsuji, Attractive quality and must-be quality, *J. Jpn. Soc. Qual. Control* 14 (2), 1984, pp. 39–48.
- [31] S.M. Kim, E. Hovy, Automatic identification of pro and con reasons in online reviews, in: *Proceedings of the COLING/ACL on Main Conference Poster Sessions*, 2006, pp. 483–490.
- [32] N. Korfiatis, M. Poulos, Using online consumer reviews as a source for demographic recommendations: a case study using online travel reviews, *Expert Syst. Appl.* 40 (14), 2013, pp. 5507–5515.
- [33] N. Korfiatis, E. García-Bariocanal, S. Sánchez-Alonso, Evaluating content quality and helpfulness of online product reviews: the interplay of review helpfulness vs. review content, *Electron. Commer. Res. Appl.* 11 (3), 2012, pp. 205–217.
- [34] D. Kuksov, Y. Xie, Pricing, frills, and customer ratings, *Mark. Sci.* 29 (5), 2010, pp. 925–943.
- [35] A.N. Langville, C.D. Meyer, Deeper inside PageRank, *Internet Math.* 1 (3), 2004, pp. 335–380.
- [36] T.Y. Lee, E.T. Bradlow, Automatic Construction of Conjoint Attributes and Levels from Online Customer Reviews. Working Paper, University of Pennsylvania, The Wharton Business School, 2007.
- [37] T.Y. Lee, E.T. Bradlow, Automated marketing research using online customer reviews, *J. Mark. Res.* 48 (5), 2011, pp. 881–894.
- [38] J. Lee, J.N. Lee, Understanding the product information inference process in electronic word-of-mouth: an objectivity–subjectivity dichotomy perspective, *Inform. Manag.* 46 (5), 2009, pp. 302–311.
- [39] F. Li, N. Liu, H. Jin, K. Zhao, Q. Yang, X. Zhu, Incorporating reviewer and product information for review rating prediction, in: *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence*, vol. 3, 2011, pp. 1820–1825.
- [40] Y.M. Li, H.M. Chen, J.H. Liou, L.F. Lin, Creating social intelligence for product portfolio design, *Decis. Support Syst.* 66, 2014, pp. 123–134.
- [41] M. McGlohon, N. Glance, Z. Reiter, Star quality: aggregating reviews to rank products and merchants, in: *Proceedings of the Fourth International AAAI Conference on Weblogs and Social Media*, 2010, pp. 114–121.
- [42] R.D. McKelvey, W. Zavoina, A statistical model for the analysis of ordinal level dependent variables, *J. Math. Soc.* 4 (1), 1975, pp. 103–120.
- [43] B. Monica, G. Marco, S. Franco, Inside PageRank, *ACM Trans. Internet Technol.* 5 (1), 2005, pp. 92–118.
- [44] S.M. Mudambi, D. Schuff, What makes a helpful online review? A study of customer reviews on amazon.com *MIS Q.* 34 (1), 2010, pp. 185–200.
- [45] A. Mukherjee, B. Liu, N. Glance, Spotting fake reviewer groups in consumer reviews, in: *Proceedings of the 21st International Conference on World Wide Web*, 2012, pp. 191–200.
- [46] O. Netzer, O. Toubia, E.T. Bradlow, E. Dahan, T. Evgeniou, F.M. Feinberg, E.M. Feit, S.K. Hui, J. Johnson, J.C. Liechty, J.B. Orlin, V.R. Rao, Beyond conjoint analysis: advances in preference measurement, *Mark. Lett.* 19 (3–4), 2008, pp. 337–354.
- [47] O. Netzer, R. Feldman, J. Goldenberg, M. Fresko, Mine your own business: market-structure surveillance through text mining, *Mark. Sci.* 31 (3), 2012, pp. 521–543.
- [48] C.W. Phang, C. Zhang, J. Sutanto, The influence of user interaction and participation in social media on the consumption intention of niche products, *Inform. Manag.* 50 (8), 2013, pp. 661–672.
- [49] A.M. Popescu, O. Etzioni, Extracting Product Features and Opinions from Reviews Chapter 2 in *Natural Language Processing and Text Mining*, Springer, 2007 pp. 9–28.
- [50] M. Porter, C. Millar, How information gives you competitive advantage, *Harv. Bus. Rev.* 63 (4), 1985, pp. 149–160.
- [51] S.M. Shugan, Commentary-relevancy is robust prediction, not alleged realism, *Mark. Sci.* 28 (5), 2009, pp. 991–998.
- [52] O. Toubia, M.G. De Jong, D. Stieger, J. Füller, Measuring consumer preferences using conjoint poker, *Mark. Sci.* 31 (1), 2012, pp. 138–156.
- [53] B. Wang, Y. Min, Y. Huang, X. Li, F. Wu, Review rating prediction based on the content and weighting strong social relation of reviewers, in: *Proceedings of the 2013 International Workshop on Mining Unstructured Big Data using Natural Language Processing*, 2013, pp. 23–30.
- [54] C.P. Wei, P. Hu, C.H. Tai, C.N. Huang, C.S. Yang, Managing word mismatch problems in information retrieval: a topic-based query expansion approach, *JMIS* 24 (3), 2007, pp. 269–295.
- [55] C.P. Wei, Y.M. Chen, C.S. Yang, C.C. Yang, Understanding what concerns consumers: a semantic approach to product feature extraction from consumer reviews, *Inform. Syst. E – Bus. Manag.* 8 (2), 2010, pp. 149–167.
- [56] Q. Xu, R.J. Jiao, X. Yang, M. Helander, H.M. Khalid, A. Opperud, An analytical Kano model for customer need analysis, *Design Stud.* 30 (1), 2009, pp. 87–110.
- [57] K. Xu, S.S. Liao, J. Li, Y. Song, Mining comparative opinions from customer reviews for competitive intelligence, *Decis. Support Syst.* 50 (4), 2011, pp. 743–754.
- [58] X. Xu, X. Cheng, S. Tan, Y. Liu, H. Shen, Aspect-level opinion mining of online customer reviews, *Commun. China* 10 (3), 2013, pp. 25–41.
- [59] Z. Yan, M. Xing, D. Zhang, B. Ma, EXPRS: an extended pagerank method for product feature extraction from online consumer reviews, *Inform. Manag.* 52 (7), 2015, pp. 850–858.
- [60] D. Yin, S. Bond, H. Zhang, Anxious or angry? Effects of discrete emotions on the perceived helpfulness of online reviews *MIS Q.* 38 (2), 2014, pp. 539–560.
- [61] Z. Zhang, C. Guo, P. Goes, Product comparison networks for competitive analysis of online word-of-mouth, *ACM Trans. Manag. Inform. Syst.* 3 (4), 2013, (article 20).
- [62] F. Zhu, X. Zhang, Impact of online consumer reviews on sales: the moderating role of product and consumer characteristics, *J. Mark.* 74 (2), 2010, pp. 133–148.

Shengsheng Xiao is Ph.D. candidate in Antai College of Economics & Management, Shanghai Jiao Tong University. He received his Master degree in Management Information Systems from Shanghai University of Finance and Economics. His research deals with trust issues in social networks, business intelligence analysis and economics of user generated content.

Chih-Ping Wei received a BS in Management Science from the National Chiao-Tung University in Taiwan, R.O.C. in 1987 and an MS and a Ph.D. in Management Information Systems from the University of Arizona in 1991 and 1996. He is currently a professor of Department of Information Management at National Taiwan University, Taiwan, R.O.C. Prior to joining National Taiwan University in 2010, he was a professor at National Tsing Hua University and National Sun Yat-sen University in Taiwan and a visiting scholar at the University of Illinois at Urbana-Champaign (Fall 2001), the Chinese University of Hong Kong (Summer 2006 and 2007), and the University of Washington (Fall 2013). His papers have appeared in *Journal of Management Information Systems (JMIS)*, *European Journal of Information Systems*, *Decision Sciences*, *Decision Support Systems*, *IEEE Transactions on Engineering Management*, *IEEE Software*, *IEEE Intelligent Systems*, *IEEE Transactions on Systems, Man, and Cybernetics*, *IEEE Transactions on Information Technology in Biomedicine*, *Journal of the American Society for Information Science and Technology*, *Information Processing and Management*, *Journal of Database Management*, and *Journal of Organizational Computing and Electronic Commerce*, etc. His current research interests include data analytics, text mining and information retrieval, knowledge management, patent analysis and intelligence, and health informatics. He can be reached at the Department of Information Management, National Taiwan University, Taipei, Taiwan, R.O.C; cpwei@im.ntu.edu.tw.

Ming Dong is Professor in the Department of Operations Management, Antai College of Economics & Management, Shanghai Jiao Tong University. He received the M.S. and Ph.D. degrees from Tianjin University, Tianjin, China, respectively, all in mechanical engineering, and the Ph.D. degree in industrial engineering from Virginia Polytechnic Institute and State University. His research and teaching interests are in the areas of decision support, operations management and production optimization.