



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

Download English Version:

<https://daneshyari.com/en/article/553748>

Download Persian Version:

<https://daneshyari.com/article/553748>

[Daneshyari.com](https://daneshyari.com)