Decision Support Systems 57 (2014) 200–211

Contents lists available at ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

Finite mixture partial least squares for segmentation and behavioral characterization of auction bidders

Ruben Mancha [a], Mark T. Leung [b,*], Jan Clark [c], Minghe Sun [b]

[a] Department of Finance and Decision Sciences, School of Business, Trinity University, United States
[b] Department of Management Science and Statistics, College of Business, University of Texas at San Antonio, United States
[c] Department of Information Systems and Cyber Security, College of Business, University of Texas at San Antonio, United States

ARTICLE INFO

Article history:
Received 6 March 2013
Received in revised form 21 August 2013
Accepted 7 September 2013
Available online 17 September 2013

Keywords:
Segmentation and profiling
Finite mixture partial least squares
Data mining
Electronic commerce
Online auction behaviors

ABSTRACT

The purpose of this study is to demonstrate how to empirically segment, without *a priori* knowledge, online auction bidders using experimental data and finite mixture models. The proposed method utilizes a finite mixture partial least squares (FIMIX-PLS) approach to examine bidder behaviors and personality characteristics, evaluate bidder differences, and then segment the bidders. The empirical experiment is conducted for two different auction mechanisms — English and Vickrey. Results from both auction mechanisms indicate that FIMIX-PLS is capable of profiling and segmenting the bidders based on their individual characteristics. The *post hoc* analysis confirms the segmentation scheme and the capability of FIMIX-PLS in segmenting bidders into statistically identifiable homogeneous groups without *a priori* information of group characteristics. Such advantage is practical for online businesses dealing with increasing amount of data about their customers on a real time basis.

1. Introduction

With the advancement of technology, many traditional services have gradually migrated to the Internet. Online business transactions and exchanges have become an integral part of modern life. As one of the many consequences, buyers and sellers around the world actively utilize the Internet as an auction house for trading their goods and services. Although online auctions are increasingly prevalent, little is known about the influences of bidding behaviors and personality characteristics on the outcome of the bidding process. At the same time, service encounter and information are viewed as two attributes imperative to strategic competition and design of effective systems in service businesses [13]. Thus, many online businesses would like to gather information on a real time basis and utilize such information to pinpoint their customer services or promotion offerings.

In most previous studies, homogeneity of single-item online bidders is generally assumed [35]. The identification of customer types is becoming a priority among business intelligence providers (*e.g.*, IBM and SAP segmentation technologies, in the case of SAP linked to their new in-memory database technology HANA), and online businesses with increasing amounts of data about their customers. In this study, we demonstrate how to empirically segment, without *a priori* knowledge, online auction bidders using experimental data and finite mixture models. This mimics the reality that online businesses may not have in advance designated customer segments or predetermined bidder characteristics associated with each segment. In addition, even previously identified online customer segments can be fluid and may change dynamically over time. Hence, it is preferable to have an analytical framework not requiring previous knowledge of segments and their characteristics. Clustering, discriminant analysis, and artificial neural networks have been applied to the problem of market segmentation [26]. The requirements of these methods vary from *a priori* identification of segments to data pre-processing and pre-classification. In particular, these classification methods fail at identifying or capturing possible interactions between the salient variables (*i.e.*, the variables relevant to the segmentation in a particular market setting). Moreover, clustering and artificial neural network methods, even efficient in finding relevant segments, are not guided by a theoretical understanding of the variables nor can they offer managers meaningful insights about particular consequences.

Our study proposes finite mixture partial least squares (FIMIX-PLS [17,52]) to identify meaningful bidding behaviors and personality characteristics for evaluating bidder differences, and then segment the bidders with respect to these profile variables. Another advantage of FIMIX-PLS is that it allows the use of endogenous variables to evaluate and account for complex and inter-connected relationships among the variables considered (*i.e.*, bidders' auction behaviors and personality characteristics). Most other traditional methods for customer segmentation such as cluster and discriminant analyses not only require *a priori* knowledge, but are also restricted to relatively simpler causality

* Corresponding author at: Department of Management Science and Statistics, College of Business, University of Texas, San Antonio, TX 78249, United States. Tel.: +1 210 458 5776; fax: +1 210 458 6350.
E-mail address: mtleung@utsa.edu (M.T. Leung).

0167-9236/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.dss.2013.09.003

relationship between the dependent variable and factors (*e.g.*, how bidder characteristics affect the bidders' willingness to pay). FIMIX-PLS can capture the more intrigue relationships on how bidder characteristics affect their online experience and behavioral outcomes (*e.g.*, bidders' enjoyment level can modify their auction behaviors that, in turn, affect their willingness to pay for an item).

Under the proposed methodological framework, auction-related variables and bidder characteristics measured in the online auction process are used as the basis for analysis and segmentation. The development of such framework should allow an online auction business to profile customers, divide them in groups sharing common characteristics, and apply tailored service encounter tools to improve both customer experience and sales opportunity. In this study, empirical experiments are devised to illustrate the use of the FIMIX-PLS framework and interpretation of the outcomes. Readers should be aware that the gist of this research is to present a general method for solving a common business problem, instead of determining the exact number of heterogeneous groups among bidders in a population or identifying their specific characteristics. It is because these analytical outcomes are not static in nature and will change from one auction item to another.

Prior studies [3,11,33,40,42,56,59] have discussed the necessity of understanding biding behaviors and investigating the effect of personality characteristics on online auction outcomes. Recent research in multi-item (Yankee) and overlapping online auctions has focused on bidder characteristics to improve understanding of bidder strategies and outcomes, in terms of price paid and hedonic value of the auction experience. For example, Bapna et al. [4–8] and Lee et al. [30] developed bidder profiles in relation to bidding behavior and motivation. Hou and Elliot [21] surveyed online shoppers who had purchased goods from a variety of auction sites to determine consumer types based on demographics and to determine motivations for participation. Hou and Rego [22] questioned the homogeneity of single-item bidders and employed cluster analysis to determine whether bidders on eBay were homogeneous. Their results showed that eBay bidders could be classified as goal-driven, experiential, playful, or opportunistic. However, their results are based on a limited sample of computer CPU auctions and did not capture individual differences.

Consumer studies have explored consumption based on hedonic value (enjoyment) in contrast to the traditional utilitarian approach (willingness to pay) to economic transaction [2]. Thus, enjoyment and willingness to pay are adopted in our study to better understand customer heterogeneity and the thrill and frenzy-related outcomes observed in online auctions (see Ref. [56] for further discussion). On the other hand, there exist a handful of studies, mostly in applied economics, linking auction bidding to hedonic valuation or rating. For example, Lange et al. [29], Noussair et al. [39] and Wertenbroch and Skiera [57] examined how bidder preferences influence the bids and, subsequently, the willingness to pay on food items. In the studies conducted by Lusk et al. [32,34], responses to different hedonic scales were used to explain bidders' perception and bid values. Therefore, we constructed a similar hedonic scale and captured the bidders' hedonic value toward the auctioned item. This hedonic construct is further linked to bidders' personality and behavior through the proposed FIMIX-PLS model.

To empirically test and validate the proposed FIMIX-PLS framework, we created an online auction system for the experiment and data collection. A fully randomized experiment was designed to reduce the effect of confounding factors on the auction outcomes (*e.g.*, the effect of winning the auction on the reported enjoyment, or the effect of web skills on winning), being able to gather individual and detailed behavioral data from the participants, while still preserving the salient features of an online auction. We conducted experiments across two, the English (ascending open bid) and the Vickrey (sealed price closed bid), auction mechanisms. This study focuses on these two auction mechanisms because they are frequently deployed in most online settings. Some well-known examples are the retailing websites such as eBay.com and uBid.com, and online advertising companies such as Google. Furthermore, the literature on online auctions frequently contrasts auction mechanisms and discusses the interaction of rules and bidder personality yielding auction behaviors [3,50,56].

In the English auction, being aware of all bids, bidders call or electronically submit their incremental bids. Some auctions may have a time limit, while others simply wait until there are no more bids. The bidder offering the highest price is the winning bidder. In the Vickrey auction, bids are not incremental, and the bidders are not aware of anyone else's bids. In a first-price sealed-bid auction, the bidder with the highest sealed bid wins the auction. The second-price sealed-bid auction is similar to the first-price sealed-bid auction, except the winner will pay the price offered by the second highest bid.

The rest of the paper is organized as follows. Section 2 describes the FIMIX-PLS method. Section 3 describes the theoretical model and the online experiments. Section 4 describes the data analyses and the application of the FIMIX-PLS method for bidder segmentation. Bidder segment characterizations for the English and Vickrey auctions are presented separately in Sections 5 and 6, respectively. Further remarks and suggestions for future research are discussed in Section 7. Section 8 concludes the paper.

2. Finite mixture segmentation

FIMIX-PLS, a derivative of partial least squares structural equation modeling (PLS-SEM) [31,58], is the method employed in this study for the exploration and statistical profiling of on-line auction bidders. A major assumption associated with PLS-SEM is that the dataset originates from a single homogeneous population, or that the subjects from different segments exhibit similar attributes and behaviors. Although this underlying assumption is met in some research scenarios, PLS-SEM is not an appropriate method when examining a diverse pool of subjects across a spectrum of background characteristics. Failing to identify the possible heterogeneity and to account for segmentation in the global PLS-SEM model may lead to inappropriate results and flawed conclusions [46,54]. Hutchinson et al. [24] discussed the basic problems arising from unobserved heterogeneity in behavioral research and outlined suggestions for this issue. As a result, several approaches, including FIMIX-PLS, have been developed to capture the heterogeneity in un-segmented datasets. Sarstedt [51] provided a survey and qualitative comparison of these approaches; and Becker et al. [9] and Ringle et al. [45] support Sarstedt's findings with simulation studies. Essentially, FIMIX-PLS allows model parameters to be estimated and subject affiliations to be simultaneously segmented [53]. Therefore, a major advantage of this approach is that it does not require *a priori* knowledge or assumptions regarding the bidder segments as seen in former studies utilizing cluster analysis (*e.g.*, [8]), neural networks (*e.g.*, [23]), and genetic algorithms (*e.g.*, [25]). Recently, Becker et al. [9] and Ringle et al. [45] examined various numerical properties of alternative PLS-SEM segmentation methods via simulation. Their findings indicated that FIMIX-PLS performs generally well relative to other methods and justified its use for segmentation. Other business applications of this multi-group analysis can be seen in Refs. [37,38,48].

For a model with finite mixture, it is assumed that the dataset is derived from a segmented population comprised of a set of subgroups that can be modeled separately. In other words, the population of online bidders can be divided into a number of segments, each with its own density function describing the characteristics and behaviors of the participants. Nevertheless, the optimal number of segments, which is not observable *a priori*, has to be determined by the FIMIX-PLS during the segmentation process. Essentially, the FIMIX-PLS approach fits well in the context of this study because there is neither general consensus as to the exact number of bidder types in the English and Vickrey online auctions, nor established literature with a structured list of attributes to segment/classify bidders. Besides, FIMIX-PLS can eliminate an additional layer of potential errors