Offering Online Recommendations to Impatient, First-Time Customers with Conjoint Based Segmentation Trees

Arnaud De Bruyn[♣]
The Pennsylvania State University

John C. Liechty
The Pennsylvania State University

Eelko K.R.E. Huizingh
University of Groningen

Gary L. Lilien
The Pennsylvania State University

Paper submitted to the Marketing Research track

European Marketing Academy 2003 Conference

The authors gratefully acknowledge the Institute for the Study of Business Markets at the Pennsylvania State University for its financial support

[♣] Contact author. 701 Business Administration Building, Smeal College of Business Administration, The Pennsylvania State University, University Park, PA 16802, USA.
Tel.: (+1) (814) 865-4091 ; Fax: (+1) (814) 865-3015 ; Email: adebruyn@psu.edu.

OFFERING ONLINE RECOMMENDATIONS TO IMPATIENT, FIRST-TIME CUSTOMERS WITH CONJOINT BASED SEGMENTATION TREES

Abstract: Online consumers are known to be impatient. It is therefore critical for merchant websites to offer relevant recommendations quickly and with minimal customer input to their online visitors. Recommending pertinent items to first-time visitors, however, is particularly challenging. In this paper, we show how conjoint analysis can be leveraged to design a recommender system that elicits customers' preferences based on simple demographics and product usage questions. We compare three algorithms, *cluster-classification*, *Bayesian treed regression* and *stepwise-conjoint regression* to develop an optimal sequence of questions. In an empirical study, *stepwise-conjoint regression* outperformed individual-based conjoint analysis to elicit individual preferences, both in term of efficiency and predictive accuracy.

Keywords: Recommender system; Segmentation tree; Conjoint analysis; Consumer decision support system

"Today's online consumers are more impatient than they ever have before"
E-Commerce Times

"Rookies are the most impatient visitors of all"
Real Business

INTRODUCTION

It is widely acknowledged that online consumers are impatient (De Angelis 2001; Enos 2001), and it is therefore crucial for e-commerce websites to offer relevant recommendations to their online visitors, quickly and with minimal customer burden. But most existing systems that offer recommendations to online visitors are poorly suited to make recommendations to new visitors, for whom no prior information is available, who may lack product-category expertise and who are often unwilling to engage in a time-consuming preference elicitation procedure.
Collaborative and *information filtering recommender systems* (Resnick and Varian 1997; Schafer et al. 2001) are "agents that use behavioral or preference information to filter alternatives and make suggestions to a user" (Ansari et al. 2000) and would appear on the surface to provide a solution. However, such systems require information about visitors' preferences, such as product ratings or implicit preferences inferred from browsing behaviors or purchase history, and are therefore impractical for first-time visitors.
Consumer-decision support systems (CDSS) potentially offer alternative solutions. A CDSS is "a system that connects a company to its existing or potential customers, providing support for some part of the customer decision-making process" (O'Keefe and McEachern 1998). Early developers assumed that by facilitating the exploitation of information and expanding processing

capabilities, users of CDSS's were likely to compare more alternatives, evaluate them more completely and thus make better decisions (Hoch and Schkade 1996). Most CDSS's assume that customers are willing and capable of comparing alternatives on those performance dimensions that are relevant and important to them. This assumption is questionable with complex, intangible or highly customizable products and where customers' expertise is low (Grenci and Todd 2002; Huffman and Kahn 1998); or when potential customers are not highly involved in the decision, impatient, and not willing to go through a time-consuming evaluation procedure. Under such circumstances, it can be risky for a commercial website to put the burden on its visitors, and count on their motivation –and patience– to use a CDSS to find the right product for them.

The objective in this paper is to develop and assess the performance of a method to offer personalized recommendations to online visitors under the following constraints:

1. The method should not require any prior knowledge about the consumer.
2. Consumers' inputs should be minimal.
3. Product-category expertise should not be needed to use the recommender system.

We first present three competing methods that use the results of a conjoint study conducted *ex ante* to develop an optimal sequence of questions and elicit customers' preferences. We then report the results of an empirical study in which each method was tested. We conclude by a discussion of the results and their managerial implications.

METHODOLOGY

Ansari et al. (2000) suggest that preference models used in marketing, such as *conjoint analysis*, offer good alternatives to collaborative and information filtering recommender systems when prior behavioral data about an individual is sparse. Despite significant efforts to make preference elicitation procedures quicker and more efficient (Sawtooth 1991; Toubia et al. 2002), conjoint analysis still requires considerable consumer input and is impractical for use as a recommender system. Nevertheless, conjoint analysis offers a useful way to model consumers' preferences. Hence, we explore the possibility that a conjoint study conducted on a sample of customers can be leveraged to design a recommender system capable of offering personal recommendations to online visitors with minimal inputs. We employ the following approach (See Figure 1):

1. We perform a conjoint analysis on a representative sample of individuals who, in addition to rating a set of products, are also asked to answer demographic and product usage questions.
2. We link respondents' characteristics to their preferences, and we identify the most informative demographic and product usage questions.
3. We use the results of analysis 2 to develop an optimal sequence of questions to elicit consumers' preferences.

{ Insert Figure 1 Here }

We now discuss three competing methods to link individuals' responses to their preferences in a way that can be operationalized in an online questionnaire.

Cluster-classification

A natural approach to this problem is to follow a two-step, segmentation-targeting strategy, similar to those commonly implemented in direct marketing. After conducting the conjoint analysis and estimating the individuals' preference partworths, we first cluster respondents into segments with similar needs and preferences, and then separate these clusters into the purest end nodes possible using the CART (Classification and Regression Trees) algorithm (Breiman et al. 1984). The tree structure conveyed in the solution can readily be translated into an optimal sequence of questions, and the partworths of the individuals populating each node can be averaged to estimate the most likely preferences of the customers with such profiles.

Bayesian treed regression

The idea behind Bayesian treed regression is to partition a dataset using a tree structure, like CART, but instead of computing a simple mean or proportion, to fit a different regression model at each end node (Chipman et al. 2002).

Although Chipman et al. (2002) did not develop the algorithm with conjoint analysis in mind, its application to this domain is natural. First, several researchers have proposed that segmenting and pooling "similar" individuals could improve prediction for each individual in conjoint analysis (DeSarbo et al. 2002; DeSarbo et al. 1989; DeSarbo et al. 1992; Green et al. 1993; Kamakura 1988; Ogawa 1987). Second, other researchers have suggested several methods to simultaneously estimate preference partworths and segment membership based on consumer descriptor variables (Gupta and Chintahunta 1994; Kamakura et al. 1994; Wedel and Steenkamp 1991; Wedel and Steenkamp 1989). Both approaches are naturally embedded in Bayesian treed regression: individuals are split (i.e., segmented) based on a sequence of relevant descriptor variables (e.g., demographic or product usage questions), and samples in each end node are pooled to estimate average preference partworths at the node's population level.

Stepwise-conjoint regression

While these tree methods are natural candidates for the design of an optimal set of questions (the successive splits signal what questions to ask and in what order, and the tree structure permits the system to choose the next best question based on the customer's previous answers), they may be subject to problems of overfitting. Because the population is successively divided, each decision to further split the population involves a smaller portion of the dataset, possibly leading to overfitting. In addition, data requirements grow exponentially with tree size.

To overcome the overfitting and data size problems, we propose a stepwise-conjoint regression approach that re-expresses the vector of preference partworths as a set of linear combinations of descriptor variables. The approach is stepwise in the sense that it sequentially selects the most informative variables to include and the issue becomes how to sequentially select the next best predictor. As with the CART algorithm, this goal can be achieved by testing all the possibilities, and keeping the one that leads to the highest incremental improvement.

In contrast to the optimal sequence of questions suggested by the two tree-based methods, stepwise-conjoint regression generates static questionnaires: *i.e.* each individual gets the same questions in the same order, independently of previous answers. However, each question selection is optimized on the entire dataset, which is likely to enhance the robustness of the method and reduce the risk of overfitting.

EMPIRICAL STUDY

Research design

We asked 616 graduate and undergraduate students at a large northeastern U.S. university to rate customized web pages from a university news portal. The pages were described by five attributes: weather report, university-related news, general news, business news, and value of an online coupon. Attributes had either two or three levels. We conducted the study electronically in a controlled lab setting. The study comprised four sections: (i) a first task, designed to familiarize respondents with the attributes and the software system used to collect the data; (ii) a conjoint task: respondents rated 21 web pages, displayed one at a time on the screen, on a 100-point preference scale; (iii) a self-administered questionnaire, with 99 questions concerning socio-demographics, consumption habits, likes and dislikes; and (iv) a holdout task where respondents distributed 100 points amongst 4 different news pages. All participants repeated the holdout task 5 times with different sets of pages in each replication.

Analysis

We randomly split the data into a training set (N=516) and a testing set (N=100).
For the cluster-classification algorithm, after estimating and scaling the individuals' partworths, we grouped respondents into clusters of preferences using the K-Means algorithm, and found that 5 groups worked best. We grew the tree using the CART algorithm and stopped the splitting process when we reached a minimum node size, and then pruned it back using the cost-complexity criterion (Breiman et al. 1984). The final tree contained 20 end nodes.
For the Bayesian treed algorithm (given our objective to require minimal consumer input), we set the two parameters that govern the splitting decisions to values that favor small trees (see Chipman et al. 2002). The final tree contained 23 end nodes.
The stepwise-conjoint regression did not require any parameterization, and statistical significance tests led us to stop the development of the model after the third question.
Preference partworths for the 100 individuals retained for testing, were calculated based on the analysis of the training sample. For the treed methods, individuals in the testing set navigated the estimated trees based on their answers to the demographic and product usage questions; for each question that they answered, partworths were an appropriately weighted average of the partworths in the remaining 'downward' nodes of the tree. For the stepwise method, partworths were simply a function of the questions answered.

Results

In table 1 we report the out-of-sample predictive accuracy of the three methods, namely the frequency each method correctly predicts the participant's top choice in holdout task, which consisted of dividing 100 points among four alternatives. The out-of-sample predictive accuracy of classic conjoint analysis estimates was 53.5%, an improvement of 28.5% compare to chance.

{ Insert Table 1 Here }

The stepwise-conjoint regression method, with a predictive accuracy of 55.9% achieved after only 3 questions, dominates the other methods, both in term of efficiency and out-of-sample predictive accuracy. The closest contender is Bayesian treed regression, with an average predictive accuracy of 49.5% at its end nodes. Despite conservative parameterization and one-third of the training set dedicated to overfitting diagnosis, the treed regression approach suffered from overfitting problems: its maximum predictive accuracy was achieved after only two questions, with 53.5%, but it failed to stop the splitting process and its performance gradually deteriorated afterwards. The cluster-classification method, with a predictive accuracy of 46.9% and a much longer sequence of questions fares far worse than the other two methods.

CONCLUSIONS

Online consumers tend to be impatient. It is therefore essential for merchant websites to offer relevant recommendations to their online visitors, quickly and with minimal consumer inputs. We have proposed a method by which companies can develop recommendation agents that are capable of providing high quality advice to first-time and impatient consumers. Specifically, we explored how traditional conjoint techniques (that have been used in the past to uncover product preferences of groups of customers to the benefit of marketing managers) could be leveraged to design recommendation agents without requiring extensive and detailed inputs usually necessary for this kind of models. We have tested alternative implementations of that approach and have shown that the stepwise conjoint regression method offers promise as a solution to this problem.

Figure 1 – A three-step approach to develop an optimal sequence of questions to elicit online visitors' preferences.

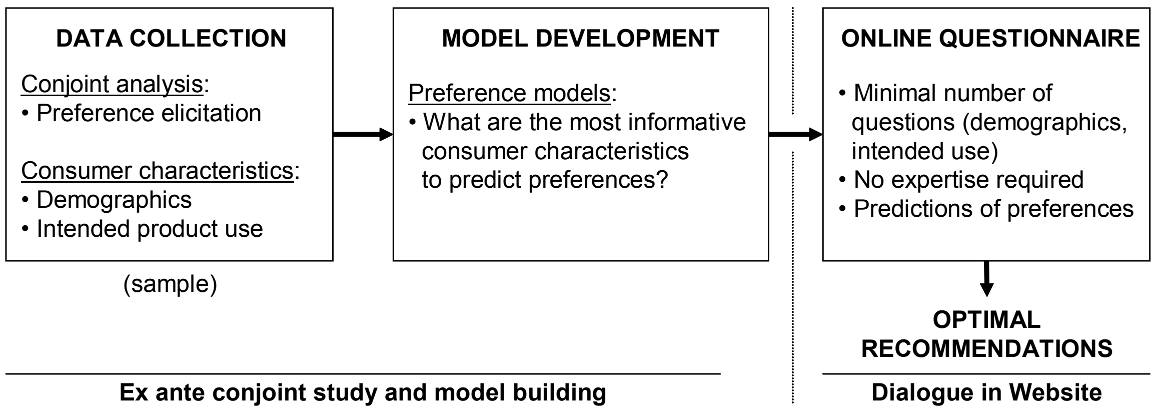

Table 1 - Comparison of the three methods plus full-profile conjoint analysis

	Full-profile conjoint	Cluster-classification	Bayesian treed regression	Stepwise-conjoint regression
Predictive accuracy	53.5%	46.9%	49.5%	55.9%
Average number of questions	21	6.2	4.2	3
Maximum number of questions	21	9	6	3
Incremental gain in predictive accuracy, per question (*)	1.4%	3.5%	5.8%	10.3%

The stepwise-conjoint regression method dominates all the other methods, including the classic, full-profile conjoint analysis, both in term of efficiency (number of questions) and out-of-sample predictive accuracy.

(*) = (predictive accuracy - 25%) / number of questions. 25% is the predictive accuracy achieved by chance.

References

Ansari, Asim, Skander Essegaier, and Rajeev Kohli (2000), "Internet Recommendation Systems," *Journal of Marketing Research*, 37 (August), 363-75.

Breiman, Leo, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone (1984), *Classification and Regression Trees*. New York: Wadsworth.

Chipman, H., E. George, and R. McCulloch (2002), "Bayesian Treed Models," *Machine Learning*, 48, 299-320.

De Angelis, Chris (2001), "Sampling for Web-Based Surveys," in *Annual Marketing Research Conference*. Atlanta, GA.

DeSarbo, Wayne S., Alexandru M. Degeratu, Michael J. Ahearne, and Kim M. Saxton (2002), "Disaggregate Market Share Response Models," *International Journal of Research in Marketing,* 19, 253-66.

DeSarbo, Wayne S., Richard L. Oliver, and Arvind Rangaswamy (1989), "A Simulated Annealing Methodology for Clusterwise Linear Regression," *Psychometrika*, 54 (December), 707-36.

DeSarbo, Wayne S., Michel Wedel, and Marco Vriens (1992), "Latent Class Metric Conjoint Analysis," *Marketing Letters*, 3 (July), 273-89.

Enos, Lori (2001), "Report: Five Keys for E-Tail Success," in E-Commerce Times.

Green, Paul E., Abba M. Krieger, and Catherine M. Schaffer (1993), "An Empirical Test of Optimal Respondent Weighting in Conjoint Analysis," *Journal of the Academy of Marketing Science*, 21 (Fall), 345-51.

Grenci, Richard T. and Peter A. Todd (2002), "Solutions-Driven Marketing," *Communications of the ACM*, 45 (3), 65-71.

Gupta, Sachin and Pradeep K. Chintahunta (1994), "On Using Demographic Variables to Determine Segment Membership in Logit Mixture Models," *Journal of Marketing Research*, 31 (Winter), 128-.

Hagerty, Michael R. (1993), "Can Segmentation Improve Predictive Accuracy in Conjoint Analysis?," *Journal of the Academy of Marketing Science*, 21 (4), 353-55.

Hoch, Stephen J. and David A. Schkade (1996), "A Psychological Approach to Decision Support Systems," *Management Science*, 42 (1), 51-64.

Huffman, C. and B. Kahn (1998), "Variety for Sale: Mass Customization or Mass Confusion?," *Journal of Retailing*, 74 (4), 491-513.

Kamakura, Wagner A. (1988), "A Least Square Procedure for Benefit Segmentation with Conjoint Experiments," *Journal of Marketing Research*, 25 (May), 157-67.

Kamakura, Wagner A., Michel Wedel, and Jagadish Agrawal (1994), "Concomitant variable latent class models for conjoint analysis," *International Journal of Research in Marketing*, 11 (5), 451-64.

Ogawa, Kohsuke (1987), "An Approach to Simultaneous Estimation and Segmentation in Conjoint Analysis," *Marketing Science*, 6 (1), 66-81.

O'Keefe, Robert M. and Tim McEachern (1998), "Web-Based Customer Decision Support Systems," *Communications of the ACM*, 41 (3), 71-78.

Resnick, Paul and Hal R. Varian (1997), "Recommender Systems," *Communications of the ACM*, 40 (3), 56-58.

Sawtooth (1991), "ACA System". Evanston, IL: Sawtooth Software.

Schafer, J. Ben, Joseph A. Konstan, and John Riedl (2001), "E-Commerce Recommendation Applications," *Data Mining and Knowledge Discovery,* 5 (1/2).

Toubia, Olivier, Duncan I. Simester, and John R. Hauser (2002), "Fast Polyhedral Adaptive Conjoint Estimation," MIT Working Paper Boston: M.I.T.

Wedel, Michel and Jan-Benedict E. M. Steenkamp (1991), "A Clusterwise Regression Method for Simultaneous Fuzzy Market Structuring and Benefit Segmentation," *Journal of Marketing Research*, 28 (November), 385-96.

---- (1989), "A Fuzzy Clusterwise Regression Approach to Benefit Segmentation," *International Journal of Research in Marketing*, 6 (March), 45-59.