

Rethinking Preference Measurement

Adriane Krapp*

Doctoral student of Marketing and Management Science, University of Hamburg

and

Henrik Sattler

Professor of Marketing and Management Science, University of Hamburg

Preferred Track:

Rethinking Market Research

* Universität Hamburg
Institut für Handel und Marketing
Adriane Krapp
Von-Melle-Park 5
20146 Hamburg
Germany
E-mail: krapp@hermes1.econ.uni-hamburg.de
Phone: +49 40 42838-2498, Fax: +49-40-42838-3650

Rethinking Preference Measurement

Abstract

Five methods for the measurement of preference are compared empirically with respect to predictive validity as well as respondents' interview evaluation considering respondents' involvement and cognitive abilities. The methods are Adaptive Conjoint Analysis (ACA), Choice Based Conjoint Analysis (CBC), two self-explicated models and a new method based on repeated stack sorting (RSS). We use preferences for pay-TV subscription offers as context for the empirical application. Data were collected from a total of 360 respondents for 3 sub-samples using an almost representative quota sampling procedure. Data analysis is in progress.

Keywords

preference measurement, conjoint analysis, self-explicated approach, predictive validity

1. Introduction

Over the past two decades, conjoint measurement has been a popular method for measuring customers' preference structures. As early as the mid 80s, Wittink and Cattin (1989) estimated the number of commercial applications at 400 per year only in the USA. In Europe, nearly 1000 conjoint studies were carried out by market research companies between 1986 and 1991 (Wittink et al., 1994). In the meantime, probably far more than 1000 conjoint analyses are performed each year in practical applications world-wide.

The popularity of conjoint measurement appears to derive, at least in part, from its presumed superiority in (predictive) validity over simpler, less expensive techniques such as self-explication approaches (Leigh et al., 1984). However, when considered in empirical studies, this superiority frequently has not been found. In some studies self-explicated models showed surprising robustness in terms of predictive validity (e.g. Green and Srinivasan, 1990; Hensel-Börner and Sattler, 2000; Srinivasan and Park, 1997; Green et al., 1993; Agarwal and Green, 1991), while other studies favour conjoint measurement (e.g. Huber et al., 1993; Green, Goldberg and Wiley, 1983; Akaah and Korgaonkar, 1983). This issue is of major practical relevance. If, at least in certain situations, the several types of conjoint measurement are not clearly superior in validity to self-explicated approaches, it becomes highly questionable whether future applications for measuring customers' preferences should be done by conjoint measurement, as self-explicated approaches have considerable advantages over conjoint measurement. There are advantages in terms of ease of data collection, data analysis and research design as well as with respect to savings of time and costs in data collection and data analysis.

Because of the inconclusive results of past research concerning the (predictive) validity of conjoint measurement compared to self-explicated approaches further research is needed (Srinivasan and Park, 1997; Hensel-Börner and Sattler, 2000). To address this issue we designed an empirical study which compares the predictive validity of the most common types of conjoint measurement (i.e. Adaptive Conjoint Analysis (ACA) and Choice Based Conjoint Analysis (CBC)) against a quite often used self-explicated model and a new self-explicated approach which implements an extremely easy evaluation task for respondents. In

addition to that we developed a new decompositional method based on repeated stack sorting (RSS). This method has some theoretical advantages because of ease and realism of the respondents' evaluation task.

Further to the direct comparison of the degree of validity obtained through the various methods, we study the influence of the respondents' involvement in the product category. This involvement may have an indirect impact on the validity of different preference measurement methods as it influences the level of cognitive control in a decision making, cf. Kroeber-Riel and Weinberg (1996, p. 360); Felser (1997, p. 65). The cognitive control is also influenced by the cognitive ability of the respondent. However, little research has been done to investigate the impacts of involvement and cognitive ability on predictive validity. A possible explanation for the inconclusive results in terms of the validity of different preference measurement methods might be the different levels of motivation and involvement or cognitive ability of the respondents over past studies. Besides this, respondents' involvement is important in getting access to respondents and in conducting time consuming interviews.

2. Methods Compared

The five methods compared in the empirical study are ACA, CBC, "traditional" self-explicated approach, a "new" self-explicated approach, and RSS.

ACA: Adaptive Conjoint Analysis (ACA) was introduced by Johnson (1987) and is widely used both in commercial and academic market research (Wittink et al., 1994). It is a hybrid method that combines self-explicated and decompositional elements. It is unique in designing tasks that are difficult for the respondent in the sense that the two alternatives to be evaluated have a similar overall utility. Here, ACA is implemented by ranking of attribute levels (five attributes are pre-ranked), the rating of the importance weights on a 1-2-3 scale, and graded paired comparisons of partial profile concepts on a 1 to 9 scale (attributes and levels used in the empirical study are described in chapter 3). No calibration concepts are evaluated.

CBC: Discrete Choice Analysis was introduced to market research by Louviere and Woodworth (1983). The respondents are asked to choose several times among a set of alternatives rather than to state a preference. Here, Choice Based Conjoint (CBC) by Sawtooth Software is used. It is implemented through seven paper and pencil versions with twenty fixed choice tasks consisting of three concept each.

"Traditional" self-explicated approach: It is feasible to evaluate the self-explicated part of ACA without the answers to the subsequent decompositional part. Thus, there is a self-explicated method available for assessment. Obviously, the respondents' evaluation of the ACA interview is not applicable to their self-explicated parts alone.

"New" self-explicated approach: The second self-explicated method applied is new to our knowledge. First, it ranks only the attribute levels that are not assumed to be pre-ranked. Afterwards, the attributes are sorted on three stacks: important, partly important, not important, and then rated on a 7-point importance scale.

RSS: A new decompositional method is introduced with "repeated stack sorting" (RSS). Like the classic conjoint analysis, it is based on a large number of full profile stimuli. However, the required task is presumably easier for the respondent than ranking those concepts. He or she is asked to divide the concepts into two stacks, the definitely rejected concepts and all others. Next, the concepts in the latter stack are divided into those that are not eligible and those that are. Again, the eligible stack is divided into concepts that belong to the accept set and those that do not. In the last step, the concepts in the accept set are ranked. The stacks can be evaluated with Linmap, cf. Srinivasan and Shocker (1973), or exploding

data, cf. Chapman and Staelin (1982). Another alternative is to model several decision rules, e.g. the disjunctive rule for the first steps and the compensatory rule for the later steps. In this study we will compare several of these models.

3. Study Design

As context for the empirical application we use pay-TV, measuring the preferred details of subscription offers. Respondents evaluated nine attributes, each having two to four levels (table 1). The definition of the attributes and levels is inspired by the offer of premiereworld, the main pay-TV provider in Germany, but some adaptations were made (for background information on pay-TV in Germany see Clement, 2000). Potential users of pay-TV come from all social backgrounds and, therefore, permit comparison of the applicability of the methods depending on level of cognitive ability and involvement. Further, by including the attributes erotic and information programmes the effects of socially desired answers can be analysed and compared across methods.

Table 1: Attribute and levels of the pay-TV example

| monthly rate | 29,90 DM | 39,90 DM | 54,90 DM | 69,90 DM |
|-----------------------------|---|---------------------------------|---------------------------|---------------------------|
| movies | blockbusters and classics | blockbusters | classics | no movies |
| soccer | all games of the premier league and international top games | top games of the premier league | no soccer | |
| formula 1 | formula 1 from 6 camera perspectives | formula 1 | no formula 1 | |
| Other sports | boxing | golf | American football, hockey | no other sports |
| erotic | erotic programme | no erotic programme | | |
| information | news, current affairs, documentations, advice | news, current affairs | documentations, advice | no information programmes |
| children's programme | children's programme | no children's programme | | |
| music | classic, traditional, Schlager, jazz | classic, jazz | traditional, Schlager | no music |

The respondents were sampled in three distinct sub-samples according to quotas on age group, gender, and formal education level. The quotas were set to match the population of a major German city. All respondents are 18 years or older, have at least one TV in their household, and none are pay-TV rejecters. Each sub-sample consists of approximately 120 respondents. In contrast to many other studies that compare different methods of measuring preferences (e.g. Agarwal and Green, 1991; Green et al., 1993; Green and Helsen, 1989; Leigh et al., 1984; Srinivasan and Park, 1997), this sample has the advantage that it is not drawn from university level students. Since the cognitive ability of the respondents might influence the predictive validity of some methods (Albrecht, 2000, p. 188), student samples might lead to biased results.

Each interview consisted of recruitment and quota questions, evaluation of involvement items, preference measurement by either ACA (sub-sample 1, including the self-explicated part of ACA), CBC (sub-sample 2), RSS (sub-sample 3) or the new self-explicated method (sub-sample 3), interview evaluation, and several holdout tasks. The comparisons between the five methods is done with respect to (1) validity and (2) respondents' evaluation of the interview.

(1) Validity: Validity was measured in terms of convergent and predictive validity. Convergent validity refers to the convergence of the results (i.e. part-worths as well as

simulated choice behavior) between the five methods under consideration. Measurement of predictive validity is obtained through hit rates (obtained from six holdout tasks for every respondent) and mean absolute error. The hit rate measures how many of the six holdout tasks are predicted correctly by a simulation based on the estimated part-worths. It is computed for each respondent individually and is aggregated afterwards. Thus, it is not applicable for CBC unless the data is evaluated with Hierarchical Bayes or another method for obtaining individual part-worths from discrete choice data. In this study we will use Hierarchical Bayes. The mean absolute error (MAE) compares the share of choice of the holdouts with the simulated shares (first choice or probabilistic). Huber et al. (1993) state that hit rate and MAE are useful together because they measure different properties: Hit rates depend primarily on the reliability of the individual models, whereas the MAE depend on the ability of the models to provide unbiased predictions.

(2) Respondents' evaluation of the interview: The evaluation of the involvement is implemented through perceived interview length (also compared to actual interview length) and the following items, which are measured on a 1 to 7 agreement scale: "The interview so far was easy to complete." – "The interview didn't take too long." – "The interview is interesting." – "The interview is fun." – "The interview is rich in variation." A negative evaluation of the interview is assumed to be an indicator for the possible use of simplifying strategies.

We hypothesise that both validity and evaluation are influenced by a respondent's involvement in pay-TV and his or her cognitive ability. According to Felser (1997) the effort put into a task depends on the involvement. Since the various preference measurement methods are prone to simplifying effects in a varying degree (Wright, 1975), their predictive validity might depend on the respondents' involvement. Strebinger et al. (2000) showed that a low (task) involvement may have a negative effect on the predictive validity of conjoint analysis. The (product category) involvement is measured here through the ownership of a VCR, a satellite dish etc., and through the following items (on a 1 to 7 agreement scale): "The ordinary TV programme is sufficient." – "I would like to have pay-TV." – "The technology behind pay-TV is interesting." – "The TV programme is often bad." – "I'm interested in pay-TV." – "Pay-TV is a good alternative to videos." – "Pay-TV is expensive." – "I will order pay-TV in the near future." – "Pay-TV is the TV of the future."

Formal education level is used as an indicator for cognitive ability. Tscheulin and Blaimont (1993) found a negative correlation between education level and predictive validity.

4. Results

This paper presents work in progress. The interviews are just finished at the time of writing and, thus, no results are available yet. Results will be presented at the EMAC 2001 in Bergen.

References

- Agarwal, M. K. and P. E. Green, 1991. Adaptive Conjoint Analysis versus self-explicated models: Some empirical results. *International Journal of Research in Marketing* 8, 141-146.
- Akaah, I. P. and P. K. Korgaonkar, 1983. An Empirical Comparison of the Predictive Validity of Self-Explicated, Huber-Hybrid, Traditional Conjoint, and Hybrid Conjoint Models. *Journal of Marketing Research* 20, 187-197.
- Albrecht, J., 2000. *Präferenzstrukturmessung: Ein empirischer Vergleich der Conjoint-Analyse mit einer kompositionellen Methode*. Frankfurt am Main: Peter Lang.

- Chapman, R. G. and R. Staelin, 1982. Exploiting Rank Ordered Choice Set Data Within the Stochastic Utility Model. *Journal of Marketing Research* 19, 288-301.
- Clement, M., 2000. *Interaktives Fernsehen - Analyse und Prognose seiner Nutzung*. Wiesbaden: Deutscher Universitäts-Verlag.
- Felser, G., 1997. *Werbe- und Konsumentenpsychologie*. Stuttgart: Schäffer-Poeschel.
- Green, P. E., S. M. Goldberg and J. B. Wiley, 1983. A Cross-Validation Test of Hybrid Conjoint Models. *Advances in Consumer Research* 10, 147-150.
- Green, P. E. and K. Helsen, 1989. Cross-Validation Assessment of Alternatives to Individual-Level Conjoint Analysis: A Case Study. *Journal of Marketing Research* 36, 346-350.
- Green, P. E., A. M. Krieger and M. K. Agarwal, 1993. A Cross Validation Test of Four Models for Quantifying Multiattribute Preference. *Marketing Letters* 4, 369-380.
- Green, P. E. and V. Srinivasan, 1990. Conjoint Analysis in Marketing: New Developments With Implications for Research and Practice. *Journal of Marketing* 3-19.
- Hensel-Börner, S. and H. Sattler, 2000. Ein empirischer Validitätsvergleich zwischen der Customized Computerized Conjoint Analysis (CCC), der Adaptive Conjoint Analysis und Self-Explicated Verfahren. *Zeitschrift für Betriebswirtschaft* 70, 705-727.
- Huber, J., D. R. Wittink, J. A. Fiedler and R. Miller, 1993. The Effectiveness of Alternative Preference Elicitation Procedures in Predicting Choice. *Journal of Marketing Research* 30, 105-114.
- Johnson, R. M., 1987. Adaptive Conjoint Analysis. 1987 Sawtooth Software Conference Proceedings, 253-265. Ketchum, ID: Sawtooth Software.
- Kroeber-Riel, W. and P. Weinberg, 1996. *Konsumentenverhalten*. München: Franz Vahlen.
- Leigh, T. W., D. B. MacKay and J. O. Summers, 1984. Reliability and Validity of Conjoint Analysis and Self-Explicated Weights: A Comparison. *Journal of Marketing Research* 26, 456-462.
- Louviere, J. J. and G. Woodworth, 1983. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. *Journal of Marketing Research* 20, 350-367.
- Srinivasan, V. and C. S. Park, 1997. Surprising Robustness of the Self-Explicated Approach to Customer Preference Structure Measurement. *Journal of Marketing Research* 34, 286-291.
- Srinivasan, V. and A. D. Shocker, 1973. Linear Programming Techniques for Multidimensional Analysis of Preferences. *Psychometrika* 38, 337-369.
- Strebinger, A., S. Hoffmann, G. Schweiger and T. Otter, 2000. Zur Realitätsnähe der Conjoint-Analyse: Der Effekt von Präsentationsformat, Involvement und Hemisphärität auf die subjektive Beurteilung der Aufgabe durch die Auskunftsperson und die Vorhersagevalidität. *Marketing ZFP* 55-74.
- Tscheulin, D. K. and C. Blaimont, 1993. Die Abhängigkeit der Prognosegüte von Conjoint-Studien von demographischen Probanden-Charakteristika. *Zeitschrift für Betriebswirtschaft* 63, 839-847.
- Wittink, D. R. and P. Cattin, 1989. Commercial use of conjoint analysis: An update. *Journal of Marketing* 53, 91-96.
- Wittink, D. R., M. Vriens and W. Burhenne, 1994. Commercial Use of Conjoint Analysis in Europe: Results and Critical Reflections. *International Journal of Research in Marketing* 11, 41-52.
- Wright, P., 1975. Consumer Choice Strategies: Simplifying vs. Optimizing. *Journal of Marketing Research* 12, 60-67.