Food Quality and Preference 34 (2014) 70–78

Contents lists available at ScienceDirect

Food Quality and Preference

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

Stability of market segmentation with cluster analysis – A methodological approach

Henriette Müller *, Ulrich Hamm

Faculty of Organic Agricultural Sciences, University of Kassel, Germany

ARTICLE INFO

Article history:
Received 8 April 2013
Received in revised form 16 December 2013
Accepted 16 December 2013
Available online 7 January 2014

Keywords:
Market segmentation
Cluster analysis
Reliability
Validity
Stability over time
Panel survey

ABSTRACT

Market segmentation is a very popular marketing tool. In the food sector, the characteristics of different consumer attitudes and consumption habits are often used as the basis for segmentation. However, the success of a target-oriented marketing approach to selected groups of consumers depends on the results of the methodology applied. So far, relatively little attention has been paid to the reliability of the analysis used for attitude-based market segmentation, to the validity or internal stability of results or to the dynamic stability over time with regard to number, size and properties of the segments.

In our study, we used data from a panel of more than 10,000 German households. The participants were segmented using a statement battery and the application of cluster analysis. In order to ensure an internally stable cluster solution, our focus was on the analytical and technical process of decision making when clustering a large dataset. A combination of various statistical measures was applied in order to enable objective decision making in the determination of the optimal number of clusters. The dynamic stability of the resulting segments was determined by confirmatory cluster analyses using data from the same individuals in three subsequent years.

The results of the analyses show that neither the internal nor the dynamic stability of market segments should be taken for granted. Therefore, marketers face the challenge of designing segment-specific marketing strategies in a way that allows changes in consumer preferences to be integrated.

1. Introduction

Market segmentation is a very popular marketing technique and its benefits are emphasised in every textbook on marketing research. One option for the food market is segmentation based on consumer attitudes and stated consumption habits, which are measured on a given scale via the evaluation of a battery of statements. Using these data, it is possible to identify market segments according to groups of consumers that are as homogeneous as possible in their attitudes and consumption habits, and which differ from other groups. Such segments can be used to select particular markets, and these are then targeted with specific marketing measures according to their characteristics (Aaker, Kumar, & Day, 2007).

However, if market changes cannot be reliably predicted, strategies for segmentation and the identification of target groups can only be successful when the segments remain stable over time. For market segmentation based on statement batteries, the 'time' stability of statement ratings is therefore a basic requirement. Stability is defined in terms of a consistent answer to a repeated, identical question (Batista-Foguet & Saris, 1997). Stability of attitudes and of attitude-based market segmentation is usually assumed *per se* (Hoek, Gendall, & Esslemont, 1996).

Nevertheless, while there has been a significant amount of research regarding consumer attitudes and their stability, rather less attention has been paid to the reliability of the methodology applied in attitude-based market segmentation, or to the validity of results and the stability over time of the number, size and properties of segments (Wedel & Kamakura, 2000).

In this context, cluster analysis is a commonly applied procedure for segmenting customers. In cluster analysis, consumers are sorted into relatively homogenous groups according to chosen criteria, so that consumers placed together in the same group, or cluster, show more similarities with each other than with those placed in other clusters. When referring to the stability of a cluster solution, it is necessary to differentiate between 'internal' and 'dynamic' stability (Wedel & Kamakura, 2000). The internal stability, or validity, of a cluster solution describes the potential for replicating segmentation results within the same or similar dataset. It is evaluated using, for example, split samples' procedures or variations of clustering methodology (e.g., Bouguessa, Wang, & Sun, 2006; Dubes & Jain, 1979; Hennig, 2007; Kuncheva & Vetrov, 2006; Levine & Domany, 2001; Meinshausen & Bühlmann, 2010;

* Corresponding author. Address: Faculty of Organic Agricultural Sciences, University of Kassel, Nordbahnhofstrasse 1a, D-37213 Witzenhausen, Germany. Tel.: +49 5542 98 1651.
E-mail address: h.mueller@uni-kassel.de (H. Müller).

0950-3293/$ - see front matter © 2014 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.foodqual.2013.12.004

Schellinck & Fenwick, 1981; Volkovich, Barzily, & Morozensky, 2008; Wu, Chen, Xiong, & Xie, 2009). According to Lange, Roth, Braun, and Buhmann (2004) and Dolnicar and Leisch (2010), the reproducibility of a cluster solution depends mainly on whether the data contain natural or true segments, as opposed to segments that are constructed by the applied clustering method. Dolnicar (2002) examined various market segmentation studies based on cluster analyses and found that they focused mainly on the interpretability of results, while the validity of the cluster solution was usually ignored. Yet, in order to evaluate dynamic stability, internal stability is a necessary prerequisite.

Dynamic stability refers to the stability of a cluster solution over time and is based on data collected in different time periods. Research has been undertaken on the dynamic stability of market segments using various datasets and different methods (see also Blocker and Flint (2007) for an overview). Calantone and Sawyer (1978) found that benefit segments in a consumer mail panel were relatively stable over the course of 2 years, whereas the allocation of individual households to the segments tended to change. Yuspeh and Fein (1982) determined segment stability via a membership predictor using discriminant function analysis, which could not predict segment membership correctly even for core members of a segment after almost a year. Farley, Winer, and Lehmann (1987) found segments based on consumption patterns in panel data to be highly unstable over 3 years. Ailawadi, Gedenk, and Neslin (1999) analysed segment stability in choice experiments using different approaches and identified relatively stable large segments but rather unstable smaller ones. Dolnicar (2004) introduced a stepwise procedure for tracking a posteriori segments, and found both stable and unstable tourist segments based on binary data. Using latent segments' models, Fonseca and Cardoso (2007) analysed stability in segments of supermarket customers, revealing significant differences in their profiles and size after 3 years. Müller (2011) repeated a choice experiment with a different sample after 18 months and found segments to be relatively stable regarding size and properties, using a latent class analysis.

Within the multivariate techniques used for market segmentation, various forms of cluster analyses are among the most popular. However, the application of clustering methodology requires that the researcher takes several (more or less) subjective decisions (Tonks, 2009). One of the most crucial elements in the practical application of cluster analysis is the determination of the optimal number of clusters, which is most often undertaken subjectively by the researcher (Dolnicar, 2002).

As the literature regarding internal stability or validity shows, it remains difficult to determine even a single cluster solution with certainty and, as a result, very much depends on the methodology applied. Consequently, it is even more challenging to determine dynamic stability. Our research approach was therefore to establish, as far as was possible, an objective and internally stable anchor solution in order to be able to determine dynamic stability through subsequent comparative analyses. The available dataset offered considerable scope insofar as the research could be undertaken within the same sample, consisting of a large number of respondents who were given the same questionnaire over four consecutive years.

In this paper, the focus is on the analytical and technical process of decision making when applying cluster analyses to a large dataset containing a statement battery. A combination of various statistical measures is used for determining the optimal number of clusters, in order to enable objective decision making. To determine the dynamic stability of the segments established in the first year, confirmatory cluster analyses are performed on data from the same individuals in three subsequent years. Thus, in order to determine the overall stability (or validity) of market segments, it is necessary to analyse firstly, the extent to which consumers change their ratings of identical statements over time and, secondly, the degree to which such changes might alter the composition of clusters over the given period.

As a first step in the following analysis, the stability of identical statements is evaluated over a period of 4 years, and the statements are combined to factors using factor analysis. The second stage examines how distinctively market segments can be formed, and how definitely consumers can be assigned to a particular segment. Following this, the extent to which assignment to a specific segment remains stable is determined and, similarly, to what extent a regrouping of people takes place.

The results allow conclusions to be drawn about the informative value of market segments based on attitudes and consumption habits, and the usefulness of customer segmentation in the medium to long term.

2. Material and methods

The analyses are based on German household panel data from the market research institute GfK (Gesellschaft für Konsumforschung). The data consist of a battery of 68 statements on a rating scale from 1 (totally disagree) to 5 (totally agree). The statements contain information regarding attitudes toward food consumption as well as stated purchasing habits and were recorded annually from 2005 to 2008 in Germany. The annual mortality rate in the GfK household panel ranged between 21% and 29% during this period of time. Only households that participated continuously in the survey for all 4 years were included, resulting in a sample of 10,001 households. For each household, additional socio-demographic data were available.

To determine the stability of the statements over time, correlation and reliability analyses with Cronbach's α were used. The statements were combined to factors using factor analysis. An exploratory factor analysis using principal axes methodology was performed via the statistical programme SPSS 19, in order to formulate the model for the year 2005. Afterwards, confirmatory factor analyses using the programme LISREL were carried out for 2005, as well as for the years 2006–2008, and factor scores were calculated for each year. The data were transferred back to SPSS and then cluster analyses were undertaken.

2.1. Clustering methodology and internal stability

First, an exploratory cluster analysis was used to determine the optimal number of clusters in year 2005. For this, the hierarchical Ward method and partitioning *K*-Means analysis were employed. For this dataset, the potential for applying and interpreting hierarchical clustering methods was limited due to large sample size (n = 10,001). Generally, the researcher is confronted with a difficult decision when determining the optimal number of clusters because usual statistical programmes are able to calculate an unlimited number of possible cluster solutions, but scarcely provide guidance for selecting the best solution (Bacher, 2001). Whereas the visual representation of a hierarchical cluster structure based on a dendrogram can assist in the selection of the most appropriate cluster solution for smaller datasets, this is difficult with large samples. For large datasets, only statistical criteria can be considered. Based on a comparison of empirical results, an overview and assessment of 30 possible criteria for determining the optimal number of clusters was published in 1985 by Milligan & Cooper. Since then, there has been considerable research into this issue and several other methods have been developed (e.g., Albatineh & Niewiadomska-Bugaj, 2011; Ben-Hur, Elisseeff, & Guyon, 2002; Chiang & Mirkin, 2010; Fraboni & Saltstone, 1992; Krieger & Green, 1999; Krolak-Schwedt & Eckes, 1992; Milligan & Cooper, 1985;