Development of a posteriori market segments over time - a tracking procedure

Friedrich Leisch
Vienna University of Technology, Austria

Sara Dolnicar
University of Wollongong, s.dolnicar@uq.edu.au

Publication Details
This paper was originally published as: Dolnicar, S & Leisch, F, Tracking a posteriori market segments over time, CD Proceedings of the 32nd European Marketing Academy Conference (EMAC), Glasgow, Scotland, 20-23rd May 2003.
Development of a posteriori market segments over time - a tracking procedure

Abstract
Segmentation has become a standard procedure in strategic marketing. A posteriori approaches are popular among practitioners and researchers and gaining importance over a priori consumer groupings. Despite the wide variety of methodological developments in the field, little attention has been given to tracking changes of a posteriori market segments over time. This article proposes a simple tracking procedure that allows testing of a posteriori segment developments over time on the basis of identical consecutive consumer surveys. It is flexible with regard to techniques chosen at each step and – through validation findings through repetition – allows thorough insight into market structure.

Disciplines
Business | Social and Behavioral Sciences

Publication Details
This paper was originally published as: Dolnicar, S & Leisch, F, Tracking a posteriori market segments over time, CD Proceedings of the 32nd European Marketing Academy Conference (EMAC), Glasgow, Scotland, 20-23rd May 2003.
Development of *a posteriori* market segments over time  
- a tracking procedure

_Friedrich Leisch_  
_Institut für Statistik und Decision Support Systems_  
_Universität Wien_

_Sara Dolnicar*²_  
_School of Management, Marketing and Employment Relations_  
_University of Wollongong_

Possible Tracks:  
Marketing Strategy and Management,  
Marketing Research

**Acknowledgment**

This research project was supported by the Austrian Science Foundation (FWF) under Grant SFB#010 (‘Adaptive Information Systems and Modelling in Economics and Management Science’) and conducted during Sara Dolnicar’s appointment at the Institute for Tourism and Leisure Studies at the Vienna University of Economics and Business Administration.

---

¹ Corresponding author: Sara Dolnicar, School of Management, Marketing and Employment Relations, University of Wollongong, Wollongong, NSW 2522, Australia, Telephone: (61 2) 4221 3862, Fax: (61 2) 4221 4154, sara_dolnicar@uow.edu.au

² Authors should have been listed in alphabetical order, but due to the EMAC regulation that one author can only be lead author once, the order had to be changed for purely formal reasons.
Development of *a posteriori* market segments over time – a tracking procedure

Abstract

Segmentation has become a standard procedure in strategic marketing. *A posteriori* approaches are popular among practitioners and researchers and gaining importance over a priori consumer groupings. Despite the wide variety of methodological developments in the field, little attention has been given to tracking changes of *a posteriori* market segments over time.

This article proposes a simple tracking procedure that allows testing of *a posteriori* segment developments over time on the basis of identical consecutive consumer surveys. It is flexible with regard to techniques chosen at each step and – through validation findings through repetition – allows thorough insight into market structure.

Introduction

Beside the wide use of *a priori* (Mazanec, 2000) segmentation approaches, splitting individuals on the basis of predefined criteria, *a posteriori* (Mazanec, 2000) approaches have increasingly gained popularity in the past decades to derive market segments on the basis of survey information (Lilien & Rangaswamy, 1998; Middleton, 1988; Myers, 1996; Smith, 1995). *A posteriori* approaches are typically conducted once, as compared to a priori studies. One possible reason is that *a posteriori* segmentation per se is exploratory and studying multiple time periods in an exploratory way (when even the one period case is full of possible pitfalls, Baumann, 2000; Ketchen & Shook, 1996; Punj & Stewart, 1983) would further increase complexity and intransparency and decrease reliability of results. The fact, however, that tracking *a posteriori* segment is extremely uncommon, is strongly supported by the literature survey conducted by Baumann (2000): among 47 *a posteriori* segmentation publications within the field of tourism not a single study reports on investigations over time.

The managerial advantages of being able to trace *a posteriori* segment trends in the marketplace include (1) validation of single *a posteriori* segment solutions that are used to build an entire marketing plan on, (2) increased insight into the changes encountered in the marketplace, (3) provision of a sound basis for forecasting, (4) the possibility of regular evaluation whether the segment(s) targeted should be held on to or switched (5) only temporary reduction of dimensionality during the segment assignment procedure and (6) applicability to typical multi-period data in tourism (non-panel format).

This article describes a simple procedure for tracking *a posteriori* market segment developments over time (based on prior theoretical suggestions by Leisch & Dolnicar, 2002 and Dolnicar, forthcoming 2003) and illustrates the procedure for an example from tourism industry.

The tracking procedure

The tracking framework suggested is a stepwise process including six steps. (1) The anchor period has to be defined: *A posteriori* market segmentation is no deterministic concept. The typically explorative nature of *a posteriori* segmentation has to be accounted for in this concept by choosing an anchor period for analysis. The anchor period is used as starting point
of investigation. It could be any period of survey data available: using the first period allows statements about the development since this period, using the last period gives insight about how the present market situation has developed. Clearly, if data for all periods is available and the tracking framework is used for explorative *ex post* investigation only, the definition of an anchor period would not be necessary. If, however, ongoing monitoring is aimed at, the anchor period choice is inevitable. (2) An *a posteriori* market segmentation solution has to be computed: Using the guest survey data set from the anchor period, a segment solution is derived. Any method that partitions the multidimensional data set in an appropriate manner can be used at this stage. The result is a number of market segments with membership assignments of each respondents to each one of the segments. (3) Market segments have to be described: Based on the answers of the segments members to the segmentation base (Wedel & Kamakura, 1998), the groups of tourists are described in detail. Furthermore relevant background variables (demographic, socioeconomic, behavioural etc.) are studied for each segment. Using e.g. discriminant analysis at this stage can help to validate the existence of heterogeneous segments with regard to these background information. (4) Data from other periods of time are assigned to the segments identified based on the anchor period: The guest survey data from the remaining periods is matched with the segment solution derived from the anchor year. This is achieved by extracting the centroids from the anchor solution, which function as representants or prototypes for the segments. The answer patterns of the respondents from the remaining periods of time are then assigned to most similar prototypes. The result are frequency distributions of segment assignments for each period of time. (5) Distributional changes over time are tested for: The distributions of respondents over segments are compared for over the time periods applying Chi square tests based on contingency tables including year and cluster membership information. Bonferroni correction of significance values is necessary if more that two periods of time are studied. At this stage it is possible to determine whether there are any significant trends in *a posteriori* segments over time. (6) Results must be validated: Validation is of utmost importance due to the exploratory nature of the framework. By including multiple periods of time, another dimension of possible influences is included which makes it even more dangerous to base managerial recommendations on one single analysis. Repetition can be conducted with different numbers of clusters, different algorithms or different anchor years. By comparing solutions and time changes a picture emerges that allows conclusions on the reliability of findings.

This stepwise tracking framework is flexible with regard to single components:

- Any kind of data can be used that is appropriate for traditional *a posteriori* segmentation.
- The choice of the anchor period year allows multiple explorative approaches. Insights into segment structure development can thus be gained from various perspectives.
- No particular partitioning algorithm needs to be used As long as each respondent is assigned to a segment deterministically, any appropriate grouping algorithm can be chosen.
- Not mentioned in the framework is that all the background information on the respondents can also be tested for inter-segment differences. Clearly, there is additional flexibility with regard to this, as any additional descriptive information on the respondents can be used for external validation and segment characterization.

In addition to the fundamental aim of this procedure, it might as well be used to determine the optimal segment solution in the first place. For this purpose it is recommended – as in the case of validation - to perform the entire procedure a number of times, with changing anchor years, changing partitioning algorithms and changing numbers of segments. Solutions with the
highest stability would be favoured, unless distinct density segments can be identified in the data (Dolnicar & Leisch, 2001).

**Tracking the development of behavioural tourist segments in Austria**

Austrian National Guest Survey data from the summer seasons 1994 and 1997 are used to illustrate the tracking framework suggested. An activity-based segmentation is constructed with the first survey used as anchor year. The sample sizes amount to 7967 for the year 1994 and 6604 for 1997. Respondents were asked to state which leisure activities they engage in during their vacation. In the data used, “1” indicates that the activity was undertaken sometimes or often whereas a “0” indicates both the fact that a respondent states not to undertake that particular pastime or that he or she has not answered the question. Hence, the data set used is in binary format, includes 14571 respondents and answers of each respondent with regard to 22 vacation activities are used as segmentation basis.

Data from the first period of time is chosen as anchor period. A self organizing feature map (SOFM, Kohonen, 1984, for applications in the field of a posteriori segmentation of guest surveys see Mazanec, 1994 and 1999, Dolnicar, 1997) functions as partitioning algorithm. A map with six prototypes is used (3 columns, 2 rows).

**Figure 1: Segment profiles for 1994**

The resulting segment profiles are provided in Figure 1, mirroring the SOFM grid. Each profile chart characterizes one of the segments. The bars indicate average agreement of the segment members, the line shows total sample average. Segment number 1 can be described as “culture tourists”. They state to have engaged in cultural activities of various kinds more often than the average summer tourist in Austria, especially the activity level with regard to “going to concerts”, “sightseeing”, “going out in the evening”, “shopping”, “going to the theatre”, “going to museums and exhibitions” and “spending the evening at a Heurigen (this is a typical Viennese restaurant)” is above average. Segment number 2 is less distinct, showing interest in both cultural activities as well as sports, segment number 3 is clearly sports-centred and segment number 5 is the typical relaxed summer tourists that spends the days hiking and going for walks. The remaining two segments are not interpreted, as it cannot be validated which proportion of the respondents are “active in all respects” or “not interested in any
activities” as opposed to being mere answer tendencies that are concentrated in segments number 4 and number 6.

In order to investigate changes over the two years the contingency table (Table 1) is constructed. The Pearson Chi-square renders highly significant outcome at the 99.9% significance level. The major trends from 1994 to 1997 are (1) the increase of the sports segment number 3, an increase in hiking tourists (number5) and a dramatic decrease in respondents stating either that they engage in all activities or have a positive answer tendency.

Table 1: Crosstabulation of segment sizes for 1994 and 1997 (computed with SOFMs)

<table>
<thead>
<tr>
<th>Segment</th>
<th>1994</th>
<th>1997</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1672</td>
<td>1337</td>
<td>3009</td>
</tr>
<tr>
<td>2</td>
<td>1321</td>
<td>1116</td>
<td>2437</td>
</tr>
<tr>
<td>3</td>
<td>1499</td>
<td>1474</td>
<td>2973</td>
</tr>
<tr>
<td>4</td>
<td>1034</td>
<td>848</td>
<td>1882</td>
</tr>
<tr>
<td>5</td>
<td>1241</td>
<td>1233</td>
<td>2474</td>
</tr>
<tr>
<td>6</td>
<td>1200</td>
<td>596</td>
<td>1796</td>
</tr>
</tbody>
</table>

The entire process is repeated for 10 segments with a different algorithm (topology representing networks as introduced by Martinetz & Schulten, 1994, TRNs) for validation purposes. Table 2 illustrates the association between the solutions. The culture segment is divided into segment number 2 and 5 (cross-sections are pointed out in Table 2 with black frames) with the main difference between these segments being the fact that no member of segment number 2 participates in organized excursions, whereas all segment members of group number 5 do. Similarly, the sports-oriented segment is split up in segment number 3 and 6. The main discriminating activity is biking. Hiking tourists remain stable (see segment number 10) although the amount of segments was almost doubled. The same is true for segment number 4: the potential negative answer tendency segment is represented by prototype 10. The remaining segments are difficult to interpret. The cross-tabulation (highly significant Chi square test) shows that segments 2 and 6 in the SOFM solution are both split up among the four new segments. The 10-segment-solution thus backs up segments number 1, 3, 4 and 5 from the SOFM computation as well as it supports the fact that the remaining group of tourists are not easily segmented in terms of vacation activities, as no stable representation can be arrived at.

Table 2: Crosstabulation of the 6 segment SOFM and 10 segment TRN solution

<table>
<thead>
<tr>
<th>SOFM with 6 segm.</th>
<th>TRN solution with 10 segments</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>members</td>
<td>893</td>
<td>16</td>
<td>522</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% SOFM</td>
<td>62%</td>
<td>1%</td>
<td>36%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>members</td>
<td>874</td>
<td>10</td>
<td>131</td>
<td>8</td>
<td>3</td>
<td>131</td>
<td>151</td>
<td>8</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>% SOFM</td>
<td>35%</td>
<td>5%</td>
<td>3%</td>
<td>53%</td>
<td>6%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>members</td>
<td>1</td>
<td>2116</td>
<td>51</td>
<td>4</td>
<td>1495</td>
<td>69</td>
<td>12</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>% SOFM</td>
<td>56%</td>
<td>1%</td>
<td>40%</td>
<td>2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>members</td>
<td>20</td>
<td>249</td>
<td>22</td>
<td>4</td>
<td>106</td>
<td>19</td>
<td>91</td>
<td>1724</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Over consecutive survey years the results from the six-segment solution are supported: an increase in the sports activity segment and the hiking segment is revealed and the decrease of segment SOFM number 6 (TRN numbers 4, 7 and 8) is true in both analyses as well.

**Table 3: Cross-tabulation of segment sizes over time**

<table>
<thead>
<tr>
<th>Segment number</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>622</td>
<td>983</td>
<td>746</td>
<td>723</td>
<td>839</td>
<td>735</td>
<td>864</td>
<td>523</td>
<td>921</td>
<td>811</td>
</tr>
<tr>
<td>1997</td>
<td>556</td>
<td>880</td>
<td>937</td>
<td>340</td>
<td>588</td>
<td>681</td>
<td>638</td>
<td>306</td>
<td>977</td>
<td>701</td>
</tr>
</tbody>
</table>

In general, thus, the validation of the initial six segments through repetition with a different number of clusters and a different partitioning algorithm pointed out the reliable findings useful for strategic marketing: The “culture tourists”, the “sports tourists”, the “hiking tourists” and the negative answer tendencies group are revealed in a stable manner repeatedly. In addition, identical market trends resulted for these groups from different analyses. The remaining segments are not identified in a stable manner and are therefore not recommendable as target segments.

**Conclusion**

A procedure for tracing trends among *a posteriori* segments over time is suggested. The advantages for strategic marketing management include (1) validation of single *a posteriori* segment solutions that are used to build an entire marketing plan on, (2) increased insight into the changes encountered in the marketplace, (3) provision of a sound basis for forecasting, (4) the possibility of regular evaluation whether the segment(s) targeted should be held on to or switched (5) only temporary reduction of dimensionality during the segment assignment procedure (no compression of the item information) and (6) applicability to typical representatively collected multi-period data (which simultaneously leads to the largest limitation of the procedure described and illustrated: the need for identical sampling procedures over both periods of time).

**References**


