



**Lancaster University**  
MANAGEMENT SCHOOL

**Lancaster University Management School**  
**Working Paper**  
**2004/028**

**Identifying Market Segments in Consumer Markets:  
Variable Selection and Data Interpretation**

Tonks, David

The Marketing Department  
Lancaster University Management School  
Lancaster LA1 4YX  
UK

©Tonks, David

All rights reserved. Short sections of text, not to exceed  
two paragraphs, may be quoted without explicit permission,  
provided that full acknowledgement is given.

The LUMS Working Papers series can be accessed at <http://www.lums.co.uk/publications>  
LUMS home page: <http://www.lums.lancs.ac.uk/>

# Title

## Identifying Market Segments in Consumer Markets: Variable Selection and Data Interpretation

### Author & affiliations

David G. Tonks  
Department of Marketing  
University of Lancaster University

Key Words: market segmentation, variables, multivariate methods, geodemographics

### Contact details

David Tonks  
Department of Marketing  
The Management School  
University of Lancaster  
Lancaster  
LA1 4YX  
UK

Telephone: (+44) 1524 593910  
Fax: (+44) 1524 593928  
Email: [d.tonks@lancaster.ac.uk](mailto:d.tonks@lancaster.ac.uk)  
Website: <http://www.lums.lancs.ac.uk/>

### Track

Working Paper  
SIG Market Segmentation Seminar, Warwick Business School, 29<sup>th</sup> June 2004

## **Abstract**

*'Market segmentation' is a key phrase if not a central notion for marketing academics and marketing practitioners alike. For example, Hooley and Saunders (1998)<sup>i</sup> propose that '... segmentation is a logical extension of the marketing concept itself.' In use however, the notion of market segmentation can be problematic. Although science-like in its procedures and applications, market segmentation presents significant operational problems and these include the selection of segmentation variables and the interpretations of resulting data - such that the managerial and also the broader cultural associations of market segmentation might both be likened to a latter day 'bureaucratic form of Sophism' (Laufer and Paradeise 1990<sup>ii</sup>) where the impact of rhetoric in its classical sense can be very powerful.*

*The desires for order, for control and also for belonging can all be connected with the need to make sense of diversity and with the resulting quest to classify and seek associations. These are all underlying features of market segmentation as either academic reflection or as managerial deliberation and subsequent action. (Similar comments also apply to the agency of rhetoric). With the practice of market segmentation, a first step is the choice of a variable or variables for identification, description and association. Multivariate approaches may take the form of numerous variables each taken individually but there are many clustering type methodologies which seek to look for commonality across all variables simultaneously. Neal and Wurst (2001)<sup>iii</sup> give a useful summary of multivariate methods used to derive market segments and the more recent techniques include artificial neural networking, latent-class models and fuzzy clustering. The conceptual and mathematical complexity of such techniques can add to the danger of researchers falling into the trap encountered by the 'six men of Indostan, to learning much inclined' (Saxe 1887)<sup>iv</sup>. It is easy to miss or misinterpret the bigger picture.*

*The particular line of enquiry in this paper is the choice of segmentation variables for multivariate approaches and the ways in which qualitative assessments, or judgments, are then made of the segments which result. Geodemographics as used in GB is the reference point, this being a well-known generic classification for multivariate methods which input census and other datasets for the creation of 'typologies' – or consumer market segments – typically using a combination of factor analysis and then cluster analysis. MOSAIC from Experian (2004)<sup>v</sup> is a particular example of a commercial application.*

*As a precursor to appropriate segmentation variable selection, the nature of alternative variables has been understood within certain classic frameworks. Using the terminology of Wedel and Kamakura (1998)<sup>vi</sup>, geodemographics can be classified as a 'general, observable' variable (or base) for market segmentation. Various other terms have been used, such as 'general customer characteristics' or simply 'general variables'. The distinction between such 'general' as opposed to 'product related' or 'specific' variables is that the former concern the features of the individual or group and the latter concern*

responses to the 'product' in question. In addition, segmentation variables can be 'observable' or 'unobservable' and this distinction concerns whether or not the variable can be measured directly or indirectly. Wedel and Kamakura (op cit) acknowledge that their resulting four-cell matrix derives directly from that proposed 30 years ago by Frank, Massy and Wind (1972)<sup>vii</sup>. With this kind of taxonomy, geodemographics falls into the category 'general, observable' and proprietary techniques such as MOSAIC have been used extensively for marketing applications in GB – the essential idea being that the geodemographic segments are discrete, with high within-group similarity and high between-group dissimilarity. Such conditions should apply to both the segments descriptors and also to any link with product-related variables.

It must be recognised that for some purposes the choice of segmentation variables is fairly self-evident but this is not always the case. The standard criteria for selection of an independent, descriptor segmentation variable are that the choice should result in segments which are measurable, accessible, substantial, stable and actionable – in addition to the segments being discrete. However, the actual choice of variables may be guided much by convention, data availability, face validity, the need for immediate operationalisation and even the culture to which the researcher belongs. With any approach, including the multivariate approach of geodemographics, the choice of input variables to create 'general, observable' segments may be fairly arbitrary.

Clustering methods also raise the vexed question of interpreting the segments which emerge whether or not the input variables are appropriate. Here, and with respect to the segments generated by the analysis, a combination of summary measures of the original input variables plus product related profile and penetration data is normally used or 'observed' – and there is potentially a great deal of such data. Such usage or observation of the available data will require some subjective, 'qualitative' assessments which could be charged with a rhetorical dimension – no matter what the context.

This paper attempts to address critically the issues of variable choice for multivariate approaches to market segmentation and also the methods used for identification of the resulting market segments.

---

*i* Hooley, G. and J.A. Saunders (1998), *Marketing strategy and competitive positioning*, London, Prentice Hall.

*ii* Laufer, Romain and Paradeise, Catherine (1990), *Marketing Democracy: Public Opinion and Media Formation in Democratic Societies*, London; Transaction Publishers

*iii* Neal, William D. and John Wurst (2001) *Advances in market segmentation*, *Marketing Research*, 13:1, pp14-18

*iv* Saxe, John G, (1887) *Referring to the Buddhist fable of the Blind Sages and the Elephant*. See

[http://www.mcps.k12.md.us/curriculum/socialstd/grade7/india/Blind\\_elephant.html](http://www.mcps.k12.md.us/curriculum/socialstd/grade7/india/Blind_elephant.html)

<sup>v</sup> See: <http://www.experian.com/uk/index.html>

---

*vi Wedel, M. and Kamakura W.A. (1998), Market Segmentation: Conceptual and Methodological Foundations. (International Series in Quantitative Marketing) Dordrecht, Kluwer Academic*

*vii Frank R.E., Massy W.F. and Wind Y. (1972), Market Segmentation, Englewood Cliffs, NJ, Prentice Hall*