Latent Class Models for Marketing Strategies

An Application to the Italian Pharmaceutical Market

Francesca Bassi

University of Padova, Italy

Abstract. In this paper, an extension of the latent class (LC) approach is applied to analyse the Italian pharmaceutical market. This sector is characterised by a high level of competitiveness, more limited budgets than years ago and, at the same time, expensive sales and promotional activities; in this context, it is very important to understand which factors influence doctors in prescribing medicines, so as to design appropriate marketing strategies. A special adaptation of the multilevel LC model is estimated to identify market segments, that is, groups of doctors similar in their attitude towards the work of pharmaceutical representatives.

Keywords: three-way data sets, finite mixture models, market segmentation

Since the early 1980s, marketing analysts and scholars have applied latent structure and other types of finite mixture models with increasing frequency. One variant of the latter, the latent class (LC) model, has perhaps been the most popular. As finite mixture models can account for respondents' heterogeneity (Dillon & Kumar, 1994) and are, therefore, a promising instrument especially for market segmentation, it is not surprising that there is a growing number of papers appearing in the marketing literature which propose special variants of finite mixture models applied to market analysis (see, among many others, Dash, Schiffman, & Berenson, 1976; DeSarbo, Wedel, Vriens, & Ramaswamy, 1992; Dillon & Mulani, 1989; Grover & Srinivasan, 1987; Jain, Bass, & Chen, 1990; Kamakura & Russel, 1989).

Recently, LC models have gained recognition as a method of segmentation, with several advantages over traditional methods (Dias & Vermunt, 2007; Magidson & Vermunt, 2002). The LC cluster analysis is a model-based clustering procedure and, as such, is a probabilistic and more flexible alternative to K-means clustering.

In this paper, an extension of the LC class approach is applied to analyse the Italian pharmaceutical market, which is the fourth largest in Europe, behind Germany, France and the UK. In 2004, there were 241 producers in Italy employing 73,550 workers (Espicom, 2006); physicians enrolled in the Italian Order of Medical Doctors totalled 347,759, among whom, almost 50,000 were general practitioners (Mariani & Ventre, 2006). The pharmaceutical sector in Italy is characterised by a high level of competitiveness, more limited budgets than years ago and, at the same time, expensive sales and promotional activities. In this context, it is very important to understand which factors influence doctors in prescribing medicines, so as to design appropriate marketing strategies.

Some recent international literature aims at understanding the determinants of doctors', and also patients', demand for drugs: One study of the Italian market is, for example, that of Coscelli (2000). Pharmaceutical industries, in general, aim at understanding what doctors require from their products and their representatives, so as to direct investments to acquire market shares, possibly without wasting resources. Enterprise profits cannot be obtained without considering customer satisfaction and, in the case of the pharmaceutical sector, the primary customer is the general practitioner prescribing medicines.

The data at my disposal were collected in a survey on Italian general practitioners. Doctors were asked to express a judgement on various aspects regarding promotion strategies that were organised by the pharmaceutical industries with which they were in contact.

The aim of this paper is to identify groups of doctors with similar attitudes towards the work of pharmaceutical representatives and, specifically, to verify if the importance assigned to the various services offered by the pharmaceutical industries varies across practitioners, to devise appropriate marketing strategies for the different segments.

I applied LC models for multilevel data (Vermunt, 2003) to identify segments in the market. Traditional LC models assume that observations are independent, but in my case this assumption was violated, since doctors judged more than one pharmaceutical industry. Multilevel LC models make it possible to modify this assumption.

I was dealing with a three-way data set in which individuals – in this case, doctors – provided multiple ratings for multiple objects – in this case, pharmaceutical industries or, which is the same in this application, brands of drugs. The multilevel LC model can cluster doctors by explicitly taking into account their judgements expressed on the promotional work by the various brands. In the following sections, an adaptation of the multilevel LC to three-way data (Vermunt, 2007) is applied. Using Vermunt's terminology,

Methodology 2009; Vol. 5(2):40–45
DOI: 10.1027/1614-2241.5.2.40
© 2009 Hogrefe & Huber Publishers

in this study, cases are doctors, situations are the industries involved in the study and attributes are the judgements collected in the survey on the seven aspects of the promotional activity of the industries. The aim of the application is to cluster doctors.

This paper is organised as follows. Firstly, the data set is briefly described. Then, the Italian pharmaceutical market is segmented by applying multilevel LC models. Lastly, some brief concluding remarks are given.

Data

The data used in this paper were collected from 489 Italian general practitioners. On a 7-point scale, doctors expressed how important the following items were in inducing them to prescribe a drug proposed by a pharmaceutical industry: (1) attention of the industry to doctors' updating (ATT), (2) frequency and regularity of visits by pharmaceutical representatives (FRE), (3) assistance on diagnostic and therapeutic problems (ASS), (4) consideration for doctors' experience and suggestions (EXP), (5) quality of training of pharmaceutical representatives (QUA), (6) information on industry activities (INF) and (7) overall quality of information and promotional activities (PRO). Some demographic characteristics of doctors were also collected: number of years since the university degree was obtained (as a proxy of age), area of the country in which they work (North, Centre, or South of Italy), size of the city in which they work (less or more than 400,000 inhabitants) and the number of patients. Doctors were asked to judge each pharmaceutical industry with which they were in contact.

Overall, 68 industries were rated, receiving from 1 to 255 judgements from doctors (total 2,537 judgements). This result describes the Italian pharmaceutical market quite well in which a group of < 20 large and well-known industries operate together with a larger group of smaller and "local" firms. Doctors differed in the number of responses given, from 1 to 8, as well as in the number of pharmaceutical industries judged.

Market Segmentation

Segmentation methods can be classified into a priori, when the type and the number of segments are determined in advance by the researcher, and post hoc, when the type and the number of segments are determined on the basis of the results of data analyses. A priori methods include log-linear models, regression, logit and discriminant analysis. Post hoc methods include clustering, automatic interaction detection techniques and mixture models. Clustering methods are the most popular tool for descriptive segmentation; for a review, see, among many others, Arabie and Hubert (1994) and Punj and Stewart (1983).

LC analysis attempts to explain the observed association between the factors that make up a multiway contingency table (Goodman, 1974) by introducing unobservable underlying classes (clusters). Green, Carmone, and Wachspress (1976) first suggested the application of LC analysis to market segmentation; other interesting applications may be found in Kamakura and Mazzon (1991), Lehman, Moore, and Elrod (1982), and Paas, Bijmolt, and Vermunt (2007).

The LC approach to clustering is model based: the fundamental assumption is that of local independence, which states that objects in the same LC share a common joint probability distribution among the observed variables (Vermunt, 1997).

Multilevel LC Models

In standard LC models (Magidson & Vermunt, 2004), it is assumed that the model parameters are the same for all units (level-1 units). The basic idea of multilevel LC models is that some of the model parameters are allowed to differ across groups, or level-2 units. These differences can be modelled by including group dummies in the model, as done in multiple-group LC analysis (Clogg & Goodman, 1984), which amounts to using a fixed-effects approach. Alternatively, in a random-effects approach, the group-specific coefficients are assumed to come from a particular distribution, whose parameters should be estimated. Depending on whether the form of the mixing distribution is specified or not, either a parametric or a nonparametric random-effects approach is obtained.

Vermunt (2003) proposed a multilevel LC model as an extension of a random-coefficients logistic regression model (Agresti, Booth, Hobert, & Caffo, 2000), in which the dependent variable is not directly observed but is rather a latent variable with several observed indicators.

Let Y_{ijk}, $i = 1, \ldots, I$, $j = 1, \ldots, J$, $k = 1, \ldots K$, denote the response of individual or level-1 unit i within group or level-2 unit j on indicator or item k; $s_k = 1, \ldots S_k$, a particular level of item k; X_{ij}, a latent variable with T classes; t, a particular LC, $t = 1, \ldots, T$; $\underline{Y}_{ij}$, the full vector of responses of case i in group j; and $\underline{s}$, a possible response pattern.

The probability structure defining a simple LC model may be expressed as follows:

$$P(\underline{Y}_{ij} = \underline{s}) = \sum_{t=1}^{T} P(X_{ij} = t) P(\underline{Y}_{ij} = \underline{s} \mid X_{ij} = t) = \sum_{t=1}^{T} P(X_{ij} = t) \prod_{k=1}^{K} P(Y_{ijk} = s_k \mid X_{ij} = t). \tag{1}$$

As specified in equation (1), the probability of observing a particular response pattern is a weighted average of class-specific probability $P(Y_{ijk} = S_k \mid X_{ij} = t)$, the weight being the probability that unit i in group j belongs to LC t. As the local independence assumption implies, indicators Y_{ijk} are assumed to be independent, conditional on LC membership.

A multilevel LC model (Vermunt, 2003) consists of a mixture model equation for level-1 and level-2 units, where a group-level discrete latent variable is introduced so that parameters are allowed to differ across LCs of groups:

© 2009 Hogrefe & Huber Publishers

$$P(\underline{Y}_j = \underline{s}) = \sum_{m=1}^{M} \left[P(W_j = m) \prod_{i=1}^{n_j} \left[\sum_{t=1}^{T} P(X_{ij} = t \mid W_j = m) \times \prod_{k=1}^{K} P(Y_{ijk} = s_k \mid X_{ij} = t) \right] \right], \quad (2)$$

where W_j denotes the latent variable at group level, assuming value m, with $m = 1, \ldots, M$ and n_j is the size of group j.

Equation (2) is obtained with the additional assumption that n_j members' responses are independent of one another, conditional on group class membership.

A natural extension of the multilevel LC model involves including level-1 and level-2 covariates to predict membership, as an extension of the LC model with concomitant variables (Dayton & McReady, 1988).

Analysis of the Italian Pharmaceutical Market

To identify market segments, multilevel LC models were estimated. Specifically, the hierarchical mixture model to cluster three-way data sets proposed by Vermunt (2007) was estimated with the Latent Gold 4.0 software (Vermunt & Magidson, 2005). I was interested in defining clusters of doctors, called, from now on (latent) classes or groups, on the basis of responses given with regard to the various brands. In this respect, I was dealing with a three-way data set, since doctors provided multiple ratings for multiple objects, that is, seven judgements on a 7-point scale with reference to all the pharmaceutical industries with which they were in contact.

The model required for the purpose is an adaptation of the standard multilevel LC model. The basic assumption is that cases may be in a different LC, depending on situation or, more specifically, cases are clustered with respect to the probability of being in a particular LC in a certain situation. The basic idea is to treat the three ways as hierarchically nested levels and to assume that there is a mixture distribution at each of the two higher levels; that is, one at case level and the other at case-in-situation level. The model assumes that cases (doctors) belong to one of V possible groups $G_1, G_2, \ldots, G_V$ with probability π_v and $\sum_{v=1}^{V} \pi_v = 1$. Conditional on belonging to g_V, in situation r cases are assumed to belong to one of L groups $H_1, H_2, \ldots, H_L$, with probability $\vartheta_{1|v}, \vartheta_{2|v}, \ldots, \vartheta_{L|v}$, and $\sum_{l=1=1}^{L} \vartheta_{l|v} = 1$, for $v = 1, \ldots, V$.

As Table 1 shows, the LC multilevel model with four LCs of doctors (V) and four classes of judgements (L, henceforth, clusters) had the lowest value of the Bayesian information criterion (BIC) index. This was chosen as the final model for the result in terms of fit and for reasons of parameter interpretability. Similar models, with only minor differences in the values of the BIC index – for example, the model with four classes and three clusters – showed parameter estimates which did not describe market segments as well as the model with four classes and four clusters.

Table 1. Model fit (BIC index) for alternative numbers of classes and clusters

L	V	BIC	Log L	npar
1	1	57,060.757	−28,365.765	42
1	2	57,068.596	−28,365.765	43
1	3	57,076.435	−28,365.765	44
1	4	57,084.273	−28,365.765	45
1	5	57,092.112	−28,365.765	46
2	1	52,821.245	−26,214.654	50
2	2	52,685.205	−26,138.795	52
2	3	52,690.608	−26,133.658	54
2	4	52,706.235	−26,133.633	56
2	5	52,721.893	−26,133.602	58
3	1	51,542.040	−25,543.696	58
3	2	51,391.486	−25,456.661	61
3	3	51,321.967	−25,410.144	64
3	4	51,324.478	−25,399.636	67
3	5	51,336.165	−25,393.727	70
4	1	51,125.633	−25,304.138	66
4	2	50,987.352	−25,219.320	70
4	3	50,901.910	−25,160.922	74
4	4	50,886.077	−27,137.328	78
4	5	50,901.638	−25,129.431	82
5	1	51,053.695	−25,236.814	74

L: number of cluster of judgements (level-1 units).
V: number of classes of doctors (level-2 units).

I carried out the estimation procedure using a few sets of starting values to avoid local maxima. Responses to items were treated as measured on an ordinal scale. Model analysis was stopped at five clusters, since the 5th cluster was of negligible size (< 2%).[1]

Estimation results obtained with the best-fitting model are listed in Table 2. Four clusters of judgements of pharmaceutical representatives' work and four classes of doctors were identified. The lower part of the table lists average judgements on each aspect of the promotional activity in the four clusters. The same results are used in Figure 1 to describe cluster profiles to aid interpretation. Cluster 3 contains the highest judgements (> 6.3), and all aspects related to promotional activity are considered very important – differences among the aspects that are judged are negligible. Cluster 1 contains high levels of responses, although aspects are differentiated: frequency and regularity of visits, overall quality of promotional activity and quality of training representatives are rated as very important, whereas information on the industry's activity is not considered important at all. Cluster 4 contains the lowest judgements (< 4.5),

[1] The purpose of this paper is to identify market segments to devise appropriate marketing strategies. According to the international literature (see, e.g., Wedel & Kamakura, 2000), market segments should have some specific proprieties: one of these is substantiality, which regards size. Specifically, the substantiality criterion is satisfied if the targeted segments represent a large enough portion of the market to ensure the profitability of targeted marketing programs.

 © 2009 Hogrefe & Huber Publishers

Table 2. Multilevel LC model – estimation results, standard errors in brackets

		Cluster 1	Cluster 2	Cluster 3	Cluster 4
Size	Size	0.3974 (0.0174)	0.3392 (0.0180)	0.1526 (0.0142)	0.1109 (0.0111)
			Conditional probabilities		
Class 1	0.4567 (0.0562)	0.3108 (0.0294)	0.5731 (0.0325)	0.1044 (0.0592)	0.0118 (0.1077)
Class 2	0.2519 (0.0442)	0.3032 (0.0346)	0.2146 (0.0373)	0.0868 (0.0414)	0.3954 (0.0456)
Class 3	0.1826 (0.0450)	0.8232 (0.0179)	0.0751 (0.0180)	0.0825 (0.0434)	0.0164 (0.1171)
Class 4	0.1089 (0.0376)	0.2646 (0.0132)	0.0890 (0.0402)	0.6200 (0.0137)	0.0265 (0.0246)
			Mean values		
ATT		6.0508 (0.0370)	4.8734 (0.0536)	6.7374 (0.0352)	3.5998 (0.0945)
FRE		6.3415 (0.0318)	5.5552 (0.0479)	6.7926 (0.0306)	4.5093 (0.1034)
ASS		5.5921 (0.0421)	4.6082 (0.0531)	6.5131 (0.0458)	3.1826 (0.0965)
EXP		5.2615 (0.0514)	4.3842 (0.0610)	6.4113 (0.0524)	2.5688 (0.1028)
QUA		6.2912 (0.0322)	5.7185 (0.0429)	6.8118 (0.0273)	4.4104 (0.0969)
INF		4.6986 (0.0623)	3.9943 (0.0717)	6.3291 (0.0638)	2.1246 (0.0980)
PRO		6.1167 (0.0334)	4.8740 (0.0496)	6.7726 (0.0294)	3.3530 (0.0861)

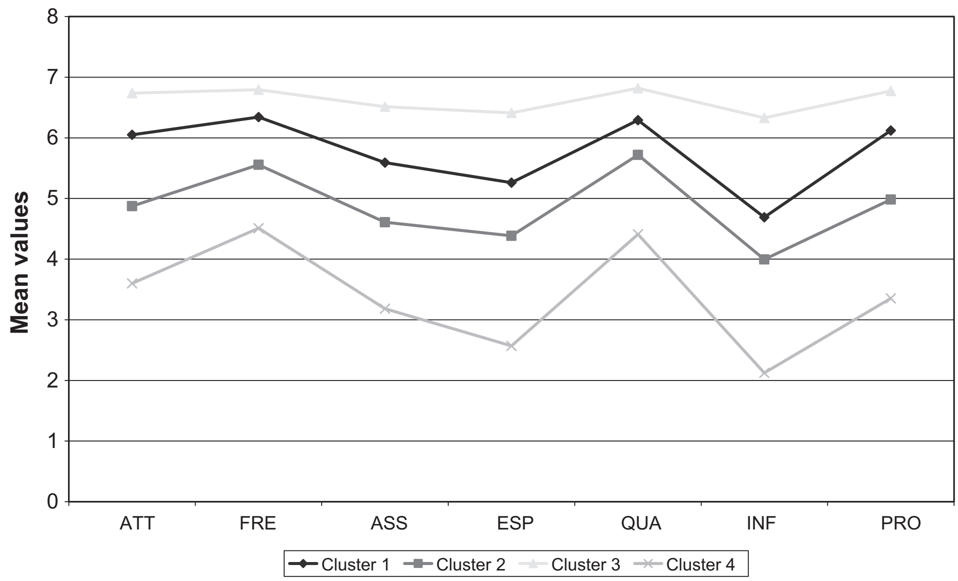

Figure 1. Cluster profiles.

meaning that no feature of promotional activity is considered important. Lastly, Cluster 2 contains judgements which fall between the other groups: some aspects, such as frequency and regularity of visits and quality of training of pharmaceutical representatives, are rated as important (average score > 5.5), the other aspects are considered almost negligible.

The upper part of Table 2 lists class sizes (π_v) and conditional probabilities $\vartheta_{l|v}$, which indicate that the four doctor-level classes have quite different distributions of judgements among clusters. Class 1 is associated with Cluster 2, Class 3 with Cluster 1 and Class 4 with Cluster 3. Class 2 contains a similar percentage of judgements assigned to Clusters 1 and 2. Looking at the results, we can try to describe doctors' segments, which is the final purpose of this analysis.

In Class 4 (11%), doctors may be defined as loyal and demanding at the same time; all items are important for them to choose among drugs. In Class 3 (18%), we find loyal practitioners who are very concerned about the frequency and regularity of visits, the quality of training of representatives, and the overall quality of information and promotional activities; information on the industry's activity is irrelevant. Class 1 (46%) contains practitioners who consider only the frequency and regularity of visits and the quality of training for representatives as important; all other aspects are not considered. Lastly, Class 2 is a mixture of two groups: a group of doctors totally unconcerned with promotion and information by industries, which prevails, and another group of practitioners only slightly interested in some aspects.

Models were first estimated without covariates. Then, doctors' demographic characteristics were included as covariates in the model with the lowest value of the BIC index. Unfortunately, they turned out not to be significant

© 2009 Hogrefe & Huber Publishers

in describing classes, which means that the groups were not significantly different with respect to doctors' age, area of the country, dimensions of the city where doctors work, and average number of patients. Relations with demographic variables usually facilitate segment identification: this was not the case in our market, where segments were defined only in terms of attitude towards promotional and information activity.

Indications from the above analyses are important for pharmaceutical industries to design appropriate promotional activities for each segment. Industries can concentrate their efforts on features of their representatives' work considered significant in each group, and not waste resources on other aspects that do not influence customers. The segment with the most interested doctors, although it was the smallest, certainly deserves great attention by representatives: conversely, industries should meditate on whether it is worth to continue to visit doctors in Class 2. A parsimonious strategy towards the remaining doctors would be to continue the frequency and regularity of visits, emphasise the quality of training and not waste resources on other aspects of promotional activity.

Instead, our segments fulfil a number of the usual criteria required for effectiveness of market segmentation (Wedel & Kamakura, 2000). Segments are large in size (substantiality); since we do not expect doctors to change their opinions on promotional activity by pharmaceutical industries frequently, segments should not change dramatically over time (stability). Segments can easily be reached by industries (accessibility) and their characteristics immediately suggest marketing strategies (actionability), revealing which aspects of promotional activity are considered most important by doctors.

Conclusive Remarks

The results presented above deserve some summarising comments in two directions: the evidence emerging about the Italian pharmaceutical market, and the models estimated.

It emerged that the Italian pharmaceutical market, at least with respect to general practitioners, is segmented. Four distinct segments of doctors can be identified, with different attitudes towards promotional activity; it is reasonable for industries to contact these four groups with diversified and appropriate strategies. One group appears very interested in all aspects considered; another group is composed of doctors not interested at all or even perhaps disturbed by promotional activity: the other two groups fall between with moderately interested doctors and others concerned only with specific aspects of pharmaceutical representatives' work.

It is also worth noting that LC models and their recent extensions deserve attention from researchers involved in market analysis. As has already been pointed out, the LC approach may help to answer questions emerging in devising marketing strategies, with reference to segmentation but not exclusively to that aspect.

Acknowledgment

The author wishes to thank Prof. Paolo Mariani for supplying the data.

References

Agresti, A., Booth, J. C., Hobert, J. P., & Caffo, B. (2000). Random-effects modeling of categorical response data. *Sociological Methodology*, *30*, 27–80.

Arabie, P., & Hubert, L. J. (1994). Cluster analysis in marketing research. In R. P. Bagozzi (Ed.), *Advanced methods in marketing research* (pp. 160–189). Cambridge, MA: Blackwell.

Clogg, C. C., & Goodman, L. A. (1984). Latent structure analysis of a set of multidimensional contingency tables. *Journal of the American Statistical Association*, *79*, 762–771.

Coscelli, A. (2000). The importance of doctors' and patients' preferences in the prescription decision. *The Journal of Industrial Economics*, *3*, 349–369.

Dash, J. F., Schiffman, L. G., & Berenson, C. (1976). Information search and store choice. *Journal of Advertising Research*, *16*, 35–40.

Dayton, M. C., & McReady, G. B. (1988). Concomitant-variable latent class models. *Journal of the American Statistical Association*, *83*, 173–178.

DeSarbo, W. S., Wedel, M., Vriens, M., & Ramaswamy, V. (1992). Latent class metric conjoint analysis. *Marketing Letters*, *3*, 273–288.

Dias, J. G., & Vermunt, J. K. (2007). Latent class modeling of website users' search patterns. Implications for online market segmentation. *Journal of Retailing and Consumer Services*, *14*, 359–368.

Dillon, W. R., & Kumar, A. (1994). Latent structure and other mixture models in marketing: An integrative survey and overview. In R. P. Bagozzi (Ed.), *Advanced methods in marketing research* (pp. 295–351). Cambridge, MA: Blackwell.

Dillon, W. R., & Mulani, N. (1989). LADI: A latent discriminant model for analysing marketing research data. *Multivariate Behavioral Research*, *19*, 438–458.

Espicom. (2006). *The outlook of pharmaceuticals: Italy*. Tangmere (UK): Espicom.

Goodman, L. A. (1974). The analysis of systems of qualitative variables when some of the variables are unobservable: Part I. A modified latent structure approach. *American Journal of Sociology*, *79*, 1179–1259.

Green, P. E., Carmone, F. J., & Wachspress, D. P. (1976). Consumer segmentation via latent class analysis. *Journal of Consumer Research*, *3*, 170–174.

Grover, R., & Srinivasan, V. (1987). A simultaneous approach to market segmentation and market structuring. *Journal of Marketing Research*, *24*, 139–153.

Jain, D. C., Bass, F. M., & Chen, Y. M. (1990). Estimation of latent class models with heterogeneous class probabilities: An application to market structuring. *Journal of Marketing Research*, *27*, 94–101.

Kamakura, W. A., & Mazzon, J. A. (1991). Value segmentation: A model for the measurement of values and value systems. *Journal of Consumer Research*, *18*, 208–218.

Kamakura, W. A., & Russel, G. J. (1989). A probabilistic choice model for market segmentation and elasticity structure. *Journal of Marketing Research*, *26*, 379–390.

Lehman, D. R., Moore, W. L., & Elrod, T. (1982). The development of distinct choice processes over time: A stochastic modeling approach. *Journal of Marketing*, *46*, 48–59.

 © 2009 Hogrefe & Huber Publishers

Magidson, J., & Vermunt, J. K. (2002). Latent class models for clustering: A comparison with K-means. *Canadian Journal of Marketing Research*, *20*, 37–44.
Magidson, J., & Vermunt, J. K. (2004). Latent class models. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 175–195). Thousand Oaks: Sage.
Mariani, P., & Ventre, G. (2006). Measure of market and territory dynamics in pharmaceutical industry: An application of shift-share technique. In *Proceedings of the XLIII SIS scientific meeting, Torino, 14–16 June* (pp. 341–344). Padova: Cleup.
Paas, L. J., Bijmolt, T. H. A., & Vermunt, J. K. (2007). Acquisition patterns of financial products: A longitudinal investigation. *Journal of Economic Psychology*, *28*, 229–241.
Punj, G. N., & Stewart, D. W. (1983). Cluster analysis in marketing research: Review and suggestions for applications. *Journal of Marketing Research*, *20*, 134–148.
Vermunt, J. K. (1997). *Loglinear models for event histories*. Thousand Oaks: Sage.
Vermunt, J. K. (2003). Multilevel latent class models. *Sociological Methodology*, *33*, 213–239.
Vermunt, J. K. (2007). A hierarchical mixture model for clustering three-way data sets. *Computational Statistics and Data Analysis*, *51*, 5368–5376.
Vermunt, J. K., & Magidson, J. (2005). *Technical guide for latent Gold 4.0: Basic and advanced*. Belmont (MA): Statistical Innovations Inc.
Wedel, M., & Kamakura, W. A. (2000). *Market segmentation: Concepts and methodological foundations*. Boston: Kluwer Academic.

Francesca Bassi

Statistics Department
University of Padova
Via C. Battisti 241
I-35121 Padova
Italy
Tel. +39 49 8274168
Fax +39 49 8274170
E-mail bassi@stat.unipd.it

© 2009 Hogrefe & Huber Publishers