FEP WORKING PAPERS

RESEARCH – WORK IN PROGRESS – N. 220, JULY 2006

EXAMINING THE SEGMENT RETENTION PROBLEM FOR THE "GROUP SATELLITE" CASE

ANA OLIVEIRA-BROCHADO

AND

F. VITORINO MARTINS

U.PORTO
FEP FACULDADE DE ECONOMIA UNIVERSIDADE DO PORTO

Faculdade de Economia da Universidade do Porto
Rua Dr. Roberto Frias, 4200-464 Porto | Tel. 225 571 100
Tel. 225571100 | www.fep.up.pt

EXAMINING THE SEGMENT RETENTION PROBLEM FOR THE "GROUP SATELLITE" CASE

Ana Oliveira-Brochado
(abrochado@fep.up.pt)

Francisco Vitorino Martins
(vmartins@fep.up.pt)

Faculdade de Economia, Universidade do Porto
Rua Dr. Roberto Frias, 4200-464 Porto, Portugal.

ABSTRACT

The purpose of this work is to determine how well, criteria designed to help the selection of the adequate number of market segments, perform in recovering small niche segments, in mixture regressions of normal data, with experimental data. The simulation experiment compares several segment retention criteria, including information criteria and classification-based criteria. We also address the impact of distributional misspecification on segment retention criteria success rates. This study shows that Akaike's Information criterion with penalty factors of 3 and 4, rather than the traditional value of 2, are the best segment retention criteria to use in recovering small niche segments. Although these criteria were designed for the specific context of mixture models, they are rarely applied in the marketing literature.

Keywords:

Information criteria; Latent Class Segmentation.

JEL-Codes: C15; C52; M31.

1 Introduction

Despite the popularity of mixture regression models in market segmentation problems, the decision about how many market segments to keep for managerial decisions is, according to many authors (DeSarbo *et al.* 1997; Wedel & DeSarbo, 1995; Wedel e Kamakura, 2000; Hawkins *et al.*, 2001; Andrews & Currim, 2003a,b), an open issue. To assess the true number of market segments is essential because many marketing decisions (segmentation, targeting, positioning, marketing mix) depended on it.

In almost all market segmentation studies, the number of market segments is determined based on heuristics as Akaike's Information Criterion (Akaike, 1973), Bayesian Information Criterion (Schwarz, 1978) and Consistent Akaike's Information Criterion[1] (Bozdogan, 1987). As the true number of market segments is unknown, it is not possible to evaluate the effectiveness of the used criteria without an experimental design.

Furthermore, a large number of new approaches were developed in the statistics literature and yet considered neither in a marketing context nor in previous studies of segment retention problem assessment (Andrews & Currim, 2003a,b, Cutler & Windham, 1994, Hawkins *et al.*, 2001).

The purpose of this work is to determine how well, criteria designed to help the selection of the adequate number of market segments, perform in what is, according to Andrews & Currim (2003) perhaps the most common analysis context in marketing research - mixture regressions of normal data. Examples of applications of these models still accumulate in the marketing literature (DeSarbo & Cron, 1988; Ramaswamy, *et al.*, 1993; Helsen *et al*, 1993; DeSarbo, *et al.*, 1992; Wedel & DeSarbo, 1994,1995; Jedidi *et al.*, 1996; DeSarbo et al., 2001; Bowman *et al.*, 2004); moreover, there are two commercial marketing research packages that accommodate mixture regression models, namely Latent Gold (the actual leader) and Glimmix (the pioneer).

---

[1] The authors reported that AIC (Akaike, 1973) was used 15 times in 37 published studies, CAIC (Bozdogan, 1987) was used 13 times and BIC (Schwarz, 1978) was used 11 times ("Determining the Number of Market Segments: an Overview", paper presented at the American Marketing Association 17th Advanced Research Techniques Forum, Montery, CA, US – June 2006).

More specifically, we intent to examine how well segment retention criteria perform in recovering both large well-separated and small weak-separated market segments into the same sample; we named this common situation in market segmentation studies as the **group satellite case**. Comparing results of previous studies, we conclude that market characteristics affect the performance of segment retention criteria, reinforcing the importance of considering this specific market condition into our experimental design.

Answering the call of a previews study (Andrews and Currim, 2003a), we also aim to evaluate the impact of distributional misspecification (when the distribution of the data do not match with the model) on segment criteria success rates. This is our second research question.

The plan of this work is as follows: we start reviewing the mixture regression model for normal, followed by a briefly description of the criteria that we aim to compare; next, we describe our experimental design used to generate the simulated data and then discuss the findings of the study.

2 Background

2.1 Multivariate Normal Mixture Regression

The latent regression model simultaneously estimate separate regression functions and memberships in S clusters

Let:

$s = 1,...,S$ indicate derived segments;
$n = 1,...,N$ indicate consumers;
$k = 1,...,K$ indicate repeated observations from individual n;
$j = 1,...,J$ indicate explanatory variables;
β_{js} = be the value of j-th regression coefficient for the s-th cluster;
Σ_s = be the covariance matrix for segment s;
y_{nk} = be the value of the dependent variable for repeated measure k on conumer s;
x_{njk} = be the value of the j-th independent variable for repeated measure k on object n.

Assume that the metric dependent vector $\mathbf{y}_n = (y_{nk})$ is distributed as a finite mixture of S conditional multivariate normal densities (1):

$$\mathbf{y}_n \sim \sum_{s=1}^{S} \lambda_s f_s\left(\mathbf{y}_n \middle| \mathbf{x}_n, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right), \tag{1}$$

where f_s is defined by the expression (2)

$$f_s\left(\mathbf{y}_n \middle| X, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right) = \left(2\pi\right)^{-K/2} \left|\Sigma_s\right|^{-1/2} \exp\left[-1/2\left(\mathbf{y}_n - \mathbf{X}\boldsymbol{\beta}_s^{'}\right)\Sigma_s^{-1}\left(\mathbf{y}_n - \mathbf{X}\boldsymbol{\beta}_s^{'}\right)'\right] \tag{2}$$

and λ_s, $s = 1,\dots,S$ are independent mixing proportions satisfying the following restrictions:

$$0 \le \lambda_s \le 1 \tag{3}$$

$$\sum_{s=1}^{S} \lambda_s = 1. \tag{4}$$

Given a sample of N independent consumers, one can thus form the likelihood (5) and the log-likelihood (6) expressions:

$$L = \prod_{n=1}^{N} \left[\sum_{s=1}^{S} \lambda_s f_s\left(\mathbf{y}_n \middle| X, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right)\right] \tag{5}$$

$$\ln L = \sum_{n=1}^{N} \ln \sum_{s=1}^{S} \lambda_s f_s\left(\mathbf{y}_n \middle| X, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right). \tag{6}$$

The implementation of the maximum likelihood procedure is done by using an Expectation-Maximization – EM type framework (Dempster *et al.*, 1977); To derive the EM algorithm is necessary to introduce non-observed data via the indicator function: $z_{ns} = 1$ if *n* comes from latent class *s* and $z_{ns} = 0$, otherwise; is assumed that z_{ns} are i.i.d multinomial. So, the joint likelihood of the "complete data" $\mathbf{y}_n = \left(y_{nk}\right)$ and $\mathbf{z}_n = \left(z_{ns}\right)$ is:

$$L_n\left(\mathbf{y}_n, \mathbf{z}_n; \mathbf{x}_n, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\lambda}\right) = \prod_{s=1}^{S} \left[\lambda_s f_s\left(\mathbf{y}_n \middle| \mathbf{x}_n, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right)\right]^{z_{ns}}. \tag{7}$$

From (7), the complete likelihood of all consumers is:

$$\ln L_c = \sum_{n=1}^{N}\sum_{s=1}^{S} z_{ns} \ln\left[f_s\left(\mathbf{y}_n \middle| \mathbf{x}_n, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right)\right] + \sum_{n=1}^{N}\sum_{s=1}^{S} z_{ns} \ln \lambda_s. \tag{8}$$

To give starting values of the parameters, the expectation (E step) and maximization (M step) of this algorithm are alternated until convergence of a sequence of log-likelihood values is obtained. Once estimates of $\boldsymbol{\lambda}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\beta}$ are obtained for any M-step, procedure, one can assign each consumer n to each latent class or market segment s via estimated posterior probability (applying Bayes' rule), providing a fuzzy clustering (E-step):

$$p_{ns} = \frac{\lambda_s f_s\left(\mathbf{y}_n \middle| \mathbf{X}, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right)}{\sum_{s=1}^{S} \lambda_s f_s\left(\mathbf{y}_n \middle| \mathbf{X}, \boldsymbol{\beta}_s, \boldsymbol{\Sigma}_s\right)}, \tag{9}$$

where $\sum_{s=1}^{S} p_{ns} = 1$, and $0 \leq p_{ns} \leq 1$.

2.2 The Criteria

We intend to compare twenty six criteria, including information criteria and classification-based criteria, described subsequently, thought a simulation experiment compares. The estimation of the order of a mixture model has been considered, mainly using a penalized form of the log-likelihood function (AIC, BIC and CAIC); as the likelihood increases with the addition of a component to a mixture model, some heuristics, called Information Criteria, attempt to balance the increase in fit obtained against the larger number of parameters estimated to models with more clusters. Information Criteria are a general family, including criteria that are estimates of (relative) Kullback-Leibler distance, approaches who been derived within a Bayesian framework for model selection and those named consistent criteria.

Although Information Criteria account for over-parameterization, as large number of clusters are derived, is also important to ensure that the segments are sufficiently separated to the selected solution. To access the ability of a mixture model, providing well-separated clusters, a statistic entropy can be used to evaluate the degree of separation in the estimated posterior probabilities. This approach yields the Classification Criteria. Some measures are derived in the context of mixture models and other are "imported" from the fuzzy literature (Bezdek *et al.*, 1997). These criteria are named respectively as probabilistic indices and fuzzy indices[2]. Table 1 presents all criteria compared in this study.

The reader is referred to the author's previous work (Brochado e Martins, 2005) and to the references cited below (Table 1) for detailed discussions of the theoretical underpinnings of these criteria.

---

[2] The quantities p_{ns} are interpreted as partial memberships in the context of fuzzy clustering and as probabilities of membership in the context of mixture models.

Table 1: Information Criteria and Classification Criteria

	Criteria	Description	Reference
INFORMATION CRITERIA		*Estimadores da distância de Kullback-Leibler*	
	Akaike Information Criteria	$\text{AIC} = -2\ln L + 2k$	Akaike (1973)
	Modified AIC 3	$\text{AIC}_3 = -2\ln L + 3k$	Bozdogan (1994)
	Modified 4	$\text{AIC}_4 = -2\ln L + 4k$	Bozdogan (1994)
	Takeuchi's Information Criterion	$\text{TIC} = -2\ln L + 2.tr\left[\mathbf{IF}^{-1}\right]$	Takeuchi (1976)
	Small sample AIC	$\text{AIC}_c = \text{AIC} + \left[2k(k+1)\right]/(N-k-1)$	Hurvich & Tsai (1989, 1995)
		Bayesian Criteria	
	Bayesian Information Criteria	$\text{BIC} = -2\ln L + k\ln N$	Schwartz (1978)
		Consistent Criteria	
	Consistent AIC	$\text{CAIC} = -2\ln L + k\left[(\ln N)+1\right]$	Bozdogan (1987)
	Information Complexity Criterion	$\text{ICOMP} = -2\ln L + k\ln\left[\frac{tr(\text{F}^{-1})}{k}\right] - \ln\left\|\mathbf{F}^{-1}\right\|$	Bozdogan (1994)
	Hannan –Quinn	$\text{HQ} = -2\ln L + 2k\ln(\ln N)$	Hannan & Quinn (1979)
	Minimum Description Length 2	$\text{MDL}_2 = -2\ln L + 2k\ln N$	Liang *et al.* (1992)
	Minimum Description Length 5	$\text{MDL}_5 = -2\ln L + 5k\ln N$	Liang *et al.* (1992)
CLASSIFICATION CRITERIA		Fuzzy Indices	
	Partition Coeficient	$\text{PC} = \sum_{n=1}^{N}\sum_{s=1}^{S} p_{ns}^2 / N$	Bezdek (1981)
	Partition Entropy	$\text{PE} = \left[\sum_{n=1}^{N}\sum_{s=1}^{S} p_{ns}\ln p_{ns}\right]/N$	Bezdek (1981)
	Normalized Partition Entropy	$\text{NPE} = \text{PE}/\left[1 - S/N\right]$	Bezdek (1981)
	Nonfuzzy Index	$\text{NFI} = \left[S\left(\sum_{n=1}^{N}\sum_{s=1}^{S} p_{ns}^2\right) - N\right]/\left[N(S-1)\right]$	Roubens (1978)
	Minimum Hard Tendency	$\text{Min}_{ht} = \max_{1\le s\le S}\left\{-\log_{10}(T_s)\right\}$	Rivera, *et al.* (1990)
	Mean Hard Tendency	$\text{Mean}_{ht} = \sum_{s=1}^{S} -\log_{10}(T_s)/S$	Rivera, *et al.* (1990)
		Probabilistic Indices	
	Entropy Measure	$\text{Es} = 1 - \left[\sum_{n=1}^{N}\sum_{s=1}^{S} -p_{ns}\ln p_{ns}\right]/N\ln S$	DeSarbo *et al.* (1992)
	Logarithm of the partition Probability	$\text{LP} = -\sum_{n=1}^{N}\sum_{s=1}^{S} z_{ns}\ln p_{ns}$	Biernacki (1997)
	Entropy	$\text{E} = -\sum_{n=1}^{N}\sum_{s=1}^{S} p_{ns}\ln p_{ns}$	Biermacki (1997)
	Normalized Entropy Criterion	$\text{NEC}(s) = \text{E}(s)/\ln L(s) - \ln L(1)$	Celeux & Soromenho (1996)
	Classification Criterion	$\text{C} = -2\ln L + 2\text{E}$	Biernacki & Govaert (1997)
	Classification Likelihood Criterion	$\text{CLC} = -2\ln L + 2\text{LP}$	Biernacki & Govaert (1997)
	Approximate Weight of Evidence	$\text{AWE} = -2\ln L_c + 2k(3/2 + \ln N)$	Banfied & Raftery (1993)
	Integrated Completed Likelihood – BIC	$\text{ICL-BIC} = -2\ln L + 2\text{LP} + k\ln N$	Biernacki & Celeux (1998)
	ICL with BIC approximation	$\text{ICOMPLBIC} = -2\ln L + 2\text{E} + k\ln N$	Dias (2004)

3 Experimental Design

3.1 The data

As our goal is to assess how segment retention criteria behave in recovering small market segments, the experiment is based on what we call the group satellite case: two large and well-separated market segments (the independent and the main group) and one small market segment, with a low degree of separation to the main group. As benchmarking case we consider two well-separated clusters with equal size. This second data enables us to evaluate in what extend segment retention criteria loose performance when we add a small market segment to the market segmentation solution. Figure 1 and Figure 2 summarize the properties of the study: number of segments, size and degree of separation between groups.

Figure 1: Satellite Group Case

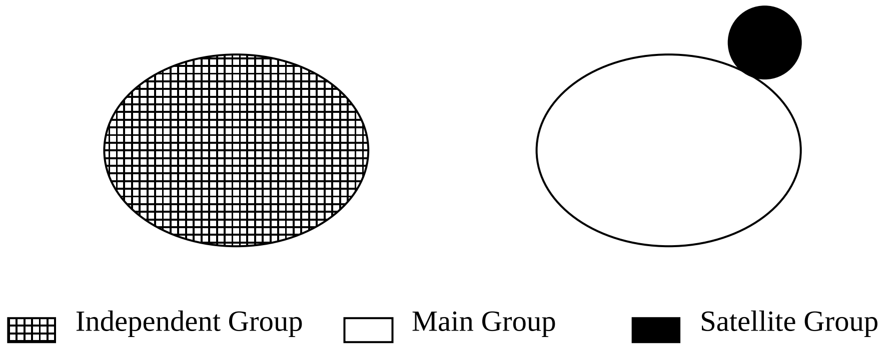
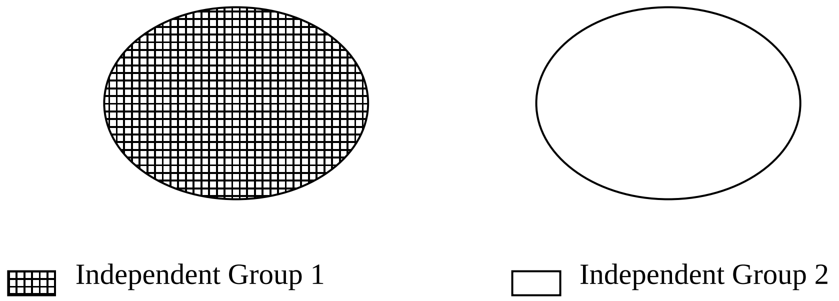


Figure 2: Benchmarking Case



In this experiment we consider six predictors, three continuous and three binary, 300 individuals with 10 observations per individual (yielding 3000 observations per data set) and an error variance of 20%. we first computed, for each subject *n* and all replications: **U**=**Xβ**; subsequently we added error to these true values **U**, $\mathbf{Y} = \mathbf{U} + \boldsymbol{\varepsilon}$; the variance of the error term was obtained from (10) (Wittink & Cattin, 1981, Vriens *et al.*, 1996)

$$PEV = \frac{\sigma_\varepsilon^2}{\sigma_\varepsilon^2 + \sigma_u^2} \Rightarrow \sigma_\varepsilon^2 = \left(\frac{PEV}{1 - PEV}\right)\sigma_u^2 \tag{10}$$

were PEV is the percent of error variance, σ_u^2 is the variance of **U** and σ_ε^2 is the variance of the error term.

In the group satellite case the mean separation between segment coefficients is set large between the Independent and the Main group (1.5) and small (0.5) between the Main and the Satellite groups. We fist randomly generate the vector of parameters β_M for the Main Group in the range of -1.5 to 1.5. Then we computed a vector with large separations $\boldsymbol{\delta}_l$ with mean 1.5 and standard deviation 0.15 and a vector with low separations $\boldsymbol{\delta}_s$ with mean 0.5 and standard deviation 0.05 of the mean; as we do not want one segment to have all coefficients larger or smaller than another because this would indicate that one segment is more sensitive that another in every way, we generate a vector of sign $\mathbf{S}_-^+$ for $\boldsymbol{\delta}_l$ and $\boldsymbol{\delta}_s$; We then compute a vector of coefficients for the Satellite Group $\boldsymbol{\beta}_{Sat} = \boldsymbol{\beta}_{Main} + \mathbf{S}_-^+\boldsymbol{\delta}_s$ (element by element) and a vector of coefficients for the Independent Group $\boldsymbol{\beta}_{Ind} = \boldsymbol{\beta}_{Main} - \mathbf{S}_-^+\boldsymbol{\delta}_l$.

Although we considered minimum segment sizes to the satellite (5% to 10%), main (40% to 50%) and independent (45%-55%) groups, the segment size is randomly generated in these ranges.

To evaluate the influence of distributional misspecification we considered two additional data experiments with an Uniform Error distribution.

To each experimental design we considered 500 experimental samples.

Table 2 summaries the experimental conditions

Table 2: Factors used in the experimental design

	Factor	Group Satellite Case	Benchmarking Case
		Level	***Level***
F1	Number of segments	3	2
F2	Number of predictors	6	6
F3	Measurememt level of predictors	Continuous & Binary	Continuous & Binary
F4	Mean separation between segment coefficients	Independent-main: large (1.5); Main-satelit: low (0.5)	Large (1.5)
F5	Number of individuals	300	300
F6	Number of observations per individual	10	10
F7	Error variance	80% (R^2=20%)	80% (R^2=20%)
F8	Segment size	5%-10% (satelit); 45%-55% (independent); 40%-50%(main)	50% (independent group 1); 50% (independent group 2)
F9	Error Distribution	Normal or Uniform	Normal or Uniform
	Number of experimental datasets	500	500

3.2 Performance Measures

We evaluate the performance of segment retention criteria by their hit rate, or percentage of datasets in which the criteria identify the correct number of segments; we also considered the over fitting rate and the under fitting rate; given two criteria with similar success rates, we prefer under fitting to over fitting; this argument is due to two empirical arguments presented in literature: first, empirical results show that over fitting produces larger parameters bias than under fitting does (Andrews & Currim, 2003a,b); second Over fitting sometimes produce very small segments with large or unstable parameter values (Cutler & Windham, 1994).

4 Results

Table 3 shows the success rates (S), rates of over fitting (O) and rates of under fitting (U) to the group satellite design and the success rates to the two benchmarking designs.

As example, to the group satellite experiment, with normal distributed errors, AIC correctly identified the true number of segments in 34% of data sets, over fitted the number of components in 48% of data sets and under fitted the number of components in 18% of these data sets. In the benchmarking case AIC has a higher success rate (68%), which is an expected result because the existence of a small market segment adds complexity to the problem.

AIC_3 and AIC_4 have the best overall performance in the simulation; although other criteria (AIC, AIC_C & CL) have an equally success rate, AIC_3 and AIC_4 exhibit lower rates of over fitting; as we previously mentioned, we prefer to avoid over fitting and, specifically for the group satellite case, a solution with two segments is much more acceptable than a solution with four segments. The criteria HQ also performed satisfactorily. BIC and CAIC – two of the most used criteria - while presenting small rates of over fitting (8% and 5%), only recovered the group satellite in 18% and 11% of the samples, respectively. As we can see on Table 3, the criteria with worst performance in this case are: MDL2, MDL5, lnL, AWE, E, LP, Es, TIC, Meanht, TIC, NEC. BIC and CAIC.

When considering the benchmarking case we conclude that not only AIC3 and AIC4 presented 100% hit rates, but also BIC, CAIC, MDL2, MDL5, HQ, AWE, ICLBIC, ICOMLPBIC; as some of these criteria (MDL2, MDL5, AWE, choose a solution with 2 segments to the group satellite case we can not conclude if they have a real good performance or a tendency to underestimate.

As expected, all criteria perform worse in the group satellite case than in the benchmarking case.

Surprisingly, segment retention criteria don't loose performance with distributional misspecification. AIC, AIC_C, AIC_3, AIC_4, BIC, CAIC and HQ exhibit similar performance rates of approximately 68%.

Table 3: Experimental results

Design	Error Distribution: Normal				Error Distribution: Uniform			
	Satelite Group			Benchmkt	Satelite Group			Benchmkt
Criteria	U	O	S	S	U	O	S	S
lnL	97%	0%	3%	92%	64%	16%	21%	0%
AIC	18%	48%	34%	68%	10%	21%	69%	98%
TIC	0%	93%	7%	0%	0%	77%	23%	10%
AIC_C	18%	48%	34%	70%	11%	21%	69%	98%
AIC_3	33%	33%	34%	100%	12%	20%	68%	100%
AIC_4	41%	26%	33%	100%	12%	20%	68%	100%
BIC	74%	8%	18%	100%	12%	20%	68%	100%
CAIC	84%	5%	11%	100%	12%	20%	68%	100%
ICOMP	28%	41%	31%	84%	7%	33%	60%	69%
MDL_2	97%	0%	3%	100%	12%	20%	68%	100%
MDL_5	100%	0%	0%	100%	19%	13%	68%	100%
HQ	44%	25%	31%	100%	12%	20%	68%	100%
ES	91%	1%	8%	88%	5%	92%	3%	0%
E	97%	0%	3%	99%	12%	88%	0%	24%
LP	96%	0%	4%	98%	12%	88%	0%	24%
AWE	99%	0%	1%	100%	12%	88%	1%	24%
NEC	92%	0%	8%	96%	12%	88%	0%	24%
CL	21%	45%	34%	44%	4%	93%	3%	2%
CLC	44%	27%	29%	61%	4%	93%	3%	2%
ICLBIC	87%	3%	10%	100%	12%	88%	1%	24%
ICOMLBIC	91%	2%	7%	100%	12%	88%	1%	24%
PC	94%	0%	6%	93%	46%	36%	18%	0%
PE	3%	78%	19%	1%	4%	91%	6%	0%
NPE	3%	78%	19%	1%	3%	92%	5%	0%
NFI	1%	71%	28%	1%	0%	55%	45%	1%
$Mean_{ht}$	94%	1%	5%	95%	100%	0%	1%	100%
Minht	52%	16%	32%	57%	65%	1%	34%	98%

5 Conclusion

In this study we aimed at exploring how segment retention criteria behave on what we named the group satellite case. Considering into the same simulated sample market segments with different degrees of separation, is an experiment condition that reflects a market condition not considered in previous studies.

We concluded that AIC_3 and AIC_4 are the best segment retention criteria to use in the group satellite case to recover the small niche segment; these criteria are modified versions of the AIC and were proposed by Bozdogan (1994) to handle with mixture data. Currently AIC_3 and AIC_4 criteria are rarely applied in the marketing literature. A previous simulation study (Andrews and Currim, 2003b) reported very good results for AIC_3 among a wide range of experimental conditions.

We also found that data distributional misspecification do not affect negatively segment retention criteria success rates; however, we intend to extend this work by considering different scenarios for distribution misspecification.

This work addresses the last major statistical deficiency of mixture regression models: the segment retention problem; to understand how segment retention criteria behave is important, as managers who make segmentation, targeting, positioning, marketing mix decisions rely on this type of heuristics to guide them on the selection of the model to pick.

References

Akaike, H. (1973), “Information Theory as an Extension of the Maximum Likelihood Principle”, In *Second International Symposium on Information Theory*, ed. Petrov e F. Csaki. Akademiai Kiado, Budapest, pp. 267-281.

Andrews, R. L. and Currim, I. S. (2003a), “A Comparison of Segment Retention Criteria for Mixture Logit Models”, *Journal of Marketing Research*, XL (2): 235-243.

Andrews, R. L., and Currim, L. S. (2003b), “Recovering and Profiling the True Segmentation Structure in Markets: an Empirical Investigation”, *International Journal of Research in Marketing*, 20: 177-192.

Banfield, J. D. e Raftery, A. E. (1993), “Model- Based Gaussian and Non-Gaussian Clustering”, *Biometrics*, 49 (3): 803-821.

Bezdek, J. C. (1981), *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum Press.

Biernacki, C. and Govaert, G. (1997), “Using the Classification Likelihood to Choose the Number of Clusters”, *Computing Science and Statistics*, 29: 451-457.

Biernacki, C., Celeux, G. and Govaert, G.. (2000), “Assessing a Mixture Model for Clustering with the Integrated Completed Likelihood”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22 (7): 719-725.

Bowman, D., Heilman, C. and Seetharaman, PB. (2004), "Determinants of Product-Use Compliance Behavior", Journal of Marketing Research, 41 (3): 324-338

Bozdogan, H. (1987), “Model Selection and Akaike´s Information Criterion (AIC): the General Theory and its Analytical Extensions”, *Psychometrika*, 52 (3): 345-370.

Bozdogan, H. (1992), “Choosing the Number of Component Clusters in the Mixture-Model Using a New Informational Complexity Criterion of the Inverse-Fisher Information Matrix”, In *Information and Classification: Concepts, Methods and Applications*, ed. O. Opitz, B. Lasen and R. Klar. Springer-Verlag, pp. 44-54.

Bozdogan, H. (1994), “Mixture-Model Cluster Analysis Using Model Selection Criteria and a New Information Measure of Complexity”, In *Proceedings of the fist US/Japan*

Conference on the Frontiers of Statistical Modelling: An Informational Approach, ed. Bozdogan, Vol. 2, Boston: Kluwer Academic Publishers, pp. 69-113.

Brochado, A. and Martins, V. (2005), Assessing the Number of Components in Mixture Models: a Review, FEP Working Paper N. º 194.

C. Biernacki (1997), *Choix de modèles en classification*. Ph.D. Thesis, Université de Technologie de Compiègne, [http://www-math.univ- comte.fr/pp_Annu/CBIERNACKI/]

Celeux, G. and Soromenho, G. (1996), "An Entropy Criterion for Assessing the Number of Clusters in a Mixture Model", *Journal of Classification*, 13 (2): 195-212.

Cutler, A. and Windham, M. P. (1994), "Information-Based Validity Functionals for Mixture Analysis". In *Proceedings of the First US/Japan Conference on the Frontiers of Statistical Modeling: An Informational Approach*, ed. H. Bozdogan, pp. 149-170.

DeSarbo, W. S. and Cron, W. L. (1988), "A Maximum Likelihhod Methodology for Clusterwise Linear Regression", *Journal of Classification*, 5: 249-282.

DeSarbo, W., Ansari, A., Chintagunta, P., Himmelberg, C., Jedidi, K., Johnson, R., Kamakura, K., Lenk, P., Srinivasan, K. and Wedel, M. (1997), "Representing Heterogeneity in Consumer Response Models; 1996 Choice Conference Participants". *Marketing Letters*, 8 (3): 335 – 348.

DeSarbo, W.S., Wedel, M., Vriens, M. and Ramaswamy, V. (1992)", Latent Class Metric Conjoint Analysis", *Marketing Letters*, 3 (3): 273-288.

DeSarbo, Wayne S., Jedidi, Kamel and Sinha, Indrajit (2001), "Customer Value Analysis in a Heterogeneous Market", *Strategic Management Journal*, 22: 845-857.

Dias, J. G. (2004), Finite Mixture Models. Review, Applications, and Computer-intensive Methods, Doctoral Dissertation, University of Groningen.

Hannan, E. J. and Quinn, B.G. (1979), "The Determination of the Order of an Autoregression", *Journal of the Royal Statistical Society*, series B 41: 190-195.

Hawkins, D. S. (1999), Determining the Number of Components in Mixtures of Linear Regression Models, Ph.D. Thesis, University of Kentucky.

Hawkins, D. S., Allen, D. M. and Stromberg, A. J. (2001), "Determining the Number of Components in Mixtures of Linear Models", *Computational Statistics and Data Analysis*, 38: 15-48.

Helsen, K., Jedidi, K. and DeSarbo, W. S. (1993), "A New Approach to Country Segmentation Utilizing Multinational Diffusion Patters", *Journal of Marketing*, 57: 60-71.

Hurvich, C. M. and Tsai, C. L. (1989), "Regression and Time Series Model Selection in Small Samples", *Biometrika*, 76: 297-307.

Jedidi, K; Ramaswamy, V.; DeSarbo, W.S. and Wedel, M. (1996), "The Disaggregate Estimation of Simultaneous Equation Models: an Application to the Price-Quality Relationship", *Journal of Structural Equation Modelling*, 3: 266-289.

Liang, Z.; Jaszczak, R. J. and Coleman, R. E. (1992), "Parameter Estimation of Finite Mixture Models Using the EM Algorithm and Information Criteria with Applications to Medical Image Processing", *IEEE Trans. Nuclear Science*, 39(4): 1126-1133.

Ramaswamy, V.; DeSarbo, W. S.; Reibstein, D. J. and Robinson, W. T. (1993), "An Empirical Pooling Approach for Estimating Marketing Mix Elasticities with PIMS Data". *Marketing Science*, 12 (1): 103-124.

Rivera, F. F., Zapata, E. L. and Carazo, J. M. (1990), "Cluster Validity Based on the Hard Tendency of the Fuzzy Classification", *Pattern Recognition Letters*, 11, 7-12.

Roubens, M. (1978), "Pattern Classification Problems and Fuzzy Sets", *Fuzzy Sets and Systems*, 1: 239-253.

Schwartz, G. (1978), "Estimating the Dimension of a Model", *The Annals of Statistics*, 6 (2): 461-464.

Takeuchi, K. (1976), "Distribution of Informational Statistics and a Criterion of Model Fitting", *Suri-Kagaku* (Mathematical Sciences), 153: 12-18 (in Japanese).

Vriens, M., Wedel, M. and Wilms, T. (1996), "Metric Conjoint Segmentation Methods: A Monte Carlo Comparison", Journal of Marketing Research, 32: 73-85.

Wedel, M. and Kamakura, W.A. (2000), Market Segmentation: Conceptual and Methodological Foundations, Kluwer Academic Publishers

Wedel, Michel and DeSarbo, Wayne S. (1994), "A Review of Recent Developments in Latent Class Regression Models", In Advanced Methods of Marketing Research, ed. Bagozzi, Richard P., Blackwell, pp. 352-388.

Wedel, Michel and DeSarbo, Wayne S. (1995), "A Mixture Likelihood Approach for Generalized Linear Models", *Journal of Classification*, 12: 21-55.

Wittink, D. R. and Cattin, D. (1981). "Alternative estimation Methods for Conjoint Analysis: a Monte Carlo Study". Journal of Marketing Research, 18: 101-106.

Recent FEP Working Papers

	Portugal: results from a large scale survey, December 2005
Nº 196	Stephen G. Donald, Natércia Fortuna and Vladas Pipiras, Local and global rank tests for multivariate varying-coefficient models, December 2005
Nº 195	Pedro Rui Mazeda Gil, The Firm's Perception of Demand Shocks and the Expected Profitability of Capital under Uncertainty, December 2005
Nº 194	Ana Oliveira-Brochado and Francisco Vitorino Martins, Assessing the Number of Components in Mixture Models: a Review., November 2005
Nº 193	Lúcia Paiva Martins de Sousa and Pedro Cosme da Costa Vieira, Um ranking das revistas científicas especializadas em economia regional e urbana, November 2005
Nº 192	António Almodovar and Maria de Fátima Brandão, Is there any progress in Economics? Some answers from the historians of economic thought, October 2005
Nº 191	Maria de Fátima Rocha and Aurora A.C. Teixeira, Crime without punishment: An update review of the determinants of cheating among university students, October 2005
Nº 190	Joao Correia-da-Silva and Carlos Hervés-Beloso, Subjective Expectations Equilibrium in Economies with Uncertain Delivery, October 2005
Nº 189	Pedro Cosme da Costa Vieira, A new economic journals' ranking that takes into account the number of pages and co-authors, October 2005
Nº 188	Argentino Pessoa, Foreign direct investment and total factor productivity in OECD countries: evidence from aggregate data, September 2005
Nº 187	Ana Teresa Tavares and Aurora A. C. Teixeira, Human Capital Intensity in Technology-Based Firms Located in Portugal: Do Foreign Multinationals Make a Difference?, August 2005
Nº 186	Jorge M. S. Valente, Beam search algorithms for the single machine total weighted tardiness scheduling problem with sequence-dependent setups, August 2005
Nº 185	Sofia Castro and João Correia-da-Silva, Past expectations as a determinant of present prices – hysteresis in a simple economy, July 2005
Nº 184	Carlos F. Alves and Victor Mendes, Institutional Investor Activism: Does the Portfolio Management Skill Matter?, July 2005
Nº 183	Filipe J. Sousa and Luís M. de Castro, Relationship significance: is it sufficiently explained?, July 2005
Nº 182	Alvaro Aguiar and Manuel M. F. Martins, Testing for Asymmetries in the Preferences of the Euro-Area Monetary Policymaker, July 2005
Nº 181	Joana Costa and Aurora A. C. Teixeira, Universities as sources of knowledge for innovation. The case of Technology Intensive Firms in Portugal, July 2005
Nº 180	Ana Margarida Oliveira Brochado and Francisco Vitorino Martins, Democracy and Economic Development: a Fuzzy Classification Approach, July 2005
Nº 179	Mário Alexandre Silva and Aurora A. C. Teixeira, A Model of the Learning Process with Local Knowledge Externalities Illustrated with an Integrated Graphical Framework, June 2005

Editor: Prof. Aurora Teixeira (ateixeira@fep.up.pt)
Download available at:
http://www.fep.up.pt/investigacao/workingpapers/workingpapers.htm
also in http://ideas.repec.org/PaperSeries.html