METRON (2016) 74:293–309
DOI 10.1007/s40300-016-0100-0

A multilevel latent class analysis of the purchasing channels among European consumers

Chiara Dal Bianco[1] · **Omar Paccagnella**[2] · **Roberta Varriale**[3]

Received: 30 November 2014 / Accepted: 15 September 2016 / Published online: 29 September 2016

Abstract This work aims at investigating similarities and differences in the ways of purchasing goods and services by European citizens—in particular the consumer behaviour on the preferred purchasing channels among web, phone, mail and sales representatives—by exploiting data collected through the Eurobarometer 69.1 survey in 2008. To this aim, we adopt a multilevel latent class solution, which allows to simultaneously cluster individuals and countries. The overall result is that most countries can be grouped in classes that follow a geographical division, while European citizens can be divided in classes with some specific profiles: a large proportion of consumers have not confidence with alternative purchasing channels yet, particularly among older respondents; most consumers still prefer to buy from sellers or providers located in their own country; more educated individuals show a widespread use of the web; a class of potential purchasers may be determined, particularly among younger people.

Keywords Consumer behaviour · Eurobarometer · European countries · Multilevel latent class · Purchasing channels · Segmentation

✉ Omar Paccagnella
omar.paccagnella@unipd.it

Chiara Dal Bianco
chiara.dalbianco@unive.it

Roberta Varriale
varriale@istat.it

1 Department of Economics, Ca’ Foscari University of Venice, 30121 Venice, Italy

2 Department of Statistical Sciences, University of Padua, 35121 Padua, Italy

3 ISTAT-Italian National Statistical Institute, 00184 Rome, Italy

1 Introduction

The growing process of globalisation, started in the 80s, has provided new motivating forces to the economic research. A modern global competition scenario, without country specific differences, was expected to become real. [22] theorised about a progressive homogenisation of consumer needs and desires, resulting in a global market in which standardised products would be commercialised. In the Levitt's theory, technological developments were supposed to have a driving role in the worldwide transformation.

Even though the studies in the second half of the 80s did not confirm the Levitt theory, it had become unquestionable that new technologies have been contributing to bring down geographical boundaries, in particular in the communication field. Despite of this, consumer needs and, consequently, the international marketing strategies of the enterprises have been influenced by cultural identities. Culture is the result of a specific national historical process and, even if local customs are exported all over the world, all the external influences are still mediated by the original culture. [35] provides a review and discussion on the role of the national culture in international marketing research.

As long as geographical boundaries become less effective, the variety of consumers, products and technological, organizational and strategic solutions widens. As a consequence, the challenging issue would be understanding the right trade-off between *standardization*, that leads to minor costs and higher quality, and *adaptation* to the consumer requests [36]. The global market, characterised by the binomial homogeneity-variety and the country specific deviations, opens up new outlooks for the international marketing and, in particular, for the international segmentation approaches.

The aim of the international segmentation is to give a structure to the existing heterogeneity among countries and/or among consumers, identifying homogeneous groups. The emerged segments could potentially show similar reactions to market stimuli. In a global perspective, the lack of an integrated and cross-national approach is an important limitation, because it is not possible to take advantage of costs reductions, economies of scales, high competition and bargaining power [21].

[17] suggest that some country characteristics may be helpful to identify differences in the diffusion patterns across countries and innovations. [36] try to exploit similarities between neighbouring countries through the implementation of a two-stage international segmentation. In the first stage, this approach uses general knowledge to group countries according to economic and cultural proximities. The most economically and strategically attractive country clusters are then selected. In the second stage, for the preferred country clusters, individual data are collected and profitable cross-national segments of consumers are selected in the individual segmentation procedure. This pan-regional strategy and similar global approaches have been widely implemented and underline the presence of a latent structure in the market.

While the two-stage international segmentation is a two-step process which considers separately the first and the second level of analysis, [5] propose an innovative model which takes into account the multilevel structure of the international segmentation issue, by studying the behaviour of consumers who are citizens of different countries. The authors implement a multilevel latent class model that simultaneously identifies country and individual level segments, to analyse the ownership of eight different financial products in 15 European countries. The segments' interpretation takes advantage of the inclusion of demographic covariates at the individual and national level. The model-based approach, the possibility of including individual and national level information, the simultaneity in the definition of the country and consumer segments are desirable features for an effective international segmentation model. Several applications have shown the potentiality of this multilevel solution in different areas, such as health, socio-economics, finance and education [3,13,14,27,28,38].

This paper aims at enriching the international segmentation literature by shedding more light on the European consumer behaviour on purchasing preferences. On the one hand, these findings could be used by policy-makers to promote and guarantee similar service levels across all European citizens. On the other hand, results could be helpful for companies to develop and implement suitable cross-country marketing strategies. Using data collected in 2008 by the Eurobarometer 69.1 survey, the purpose of this work is to investigate similarities and differences in the ways of purchasing goods and services in the last 12 months by the European citizens, adopting a multilevel latent class analysis as segmentation tool. In particular, the focus is on consumer behaviour concerning national and international preferred purchasing channels among web, phone, mail and sales representatives within the European member countries.

The rest of the paper is organised as follows. Section 2 describes the main features of the different non-standard purchasing channels. Sections 3 and 4 explain the data and the statistical solution adopted in this work, respectively. The main results and findings of the analysis are reported in Sect. 5, while Sect. 6 ends the paper with the conclusions and some suggestions for future research.

2 Purchasing channels

In the purchasing process, the choice of the channel for buying products could be placed between the goods or services options' evaluation and the final purchase. The consumer has to make a decision on which product to buy and, at the same time, on which purchasing channel to opt for.

The possible channels through which a company interacts with its customers are several and rapidly growing. According to [29], channel options fall mainly into six categories balancing *physical* and *virtual* contacts: sales force (including field account management, service, and personal representation), outlets (including retail, branches, stores, depots, and kiosks), telephony (including traditional telephone, facsimile, telex, and call center contact), direct marketing (including direct mail, radio, and traditional television), e-commerce (including e-mail, Internet, and interactive digital television) and m-commerce (including mobile-telephony, SMS and text messaging, and WAP and 3G mobile services).

The relevance and profitability of each channel strategy depend on needs and preferences of the target customer segment [7]. To fulfil consumers' needs, firms pursue specific strategies that range from a *mono-channel provider strategy*, based on one main channel, to a *customer segment channel strategy*, based on the evidence that different consumers may prefer to acquire information or purchase via different channels. Moreover, *integrated multichannel strategies* are even more comprehensive and aim at providing the consumer with various possibilities among which to choose [11,29]. Starting from the early 2000s, market research has shown that a large proportion of customers combine catalogues, bricks-and-mortar stores and Internet, and most of the customers have visited at least one additional channel before purchasing from another one [30].

In this paper, we focus on the consumer choice among Internet, telephone, mail and sales representative, covering four of the six categories of purchase channels proposed in the literature: e-commerce, telephony, direct market and sales force, respectively.

At the beginning, it is useful to distinguish between *physical* and *virtual* contact, among customers and sellers as well as among customers and products. The sales representative channel belongs to the former type of contact, while Internet, telephone and mail to the latter.

The representative, who visits the consumer at home, has the challenging role of gaining customers' confidence. The so called "foot-in-the-door technique" is the main strategy to increase the probability of placing an order [34]. By creating a familiar atmosphere and giving the opportunity of examining and testing the products, the seller appeals to a more emotionally purchasing process. However, since there is no opportunity to compare the proposed products with the commercialised equivalent ones, the perceived risk is very high and usually the consumer tends to prefer the traditional retailer.

When there is no direct contact with the product, the major risks perceived by consumers are receiving something with different features with respect to the ordered product, detecting and returning manufacturing defects or *flawed wares* and being compensated for the experienced problem.

The phone channel is often used in commercialising services. For instance, phone companies have massive information on their customers and could offer ad-hoc contracts. For real goods, phone buying is generally mediated by telesales, and this allows a virtual view of the product reducing the perceived risks explained above.

The mail purchases are the ancestors of the web ones. In the first case, consumers look through a paper catalogue, while in the second case through a virtual one. The paper catalogue may "put product information at a shopper's fingertips without necessitating computer and Internet access" [43]. When the amount of the transaction is moderate and convenient, the consumer tends to take the risk.

Web purchases are undoubtedly the most common alternative to the retail dealer and literature about e-commerce and e-shoppers' segmentation is extremely wide [4,16]. The web purchasing behaviour could be analysed according to four dimensions [8]: access conditions, technology experience, purchase knowledge and web experience. Access conditions are merely a computer ownership, a web access and netsurfing basic abilities. Technology experience depends on the primary reason for using the web: the principal implication of a lack of confidence with the virtual environment is the unwillingness of providing personal and credit card details. A web purchase also prevents a direct contact with the product, as well as a social relation with the provider, having an impact on the consumer sensory experience. This represents an additional noteworthy dimension in the purchase process and can be differently perceived by different cultures: for instance, the Mediterranean culture gives an important role to the sensory faculty, while in the Anglo-Saxon area this is not so relevant. Finally, web experience is directly correlated with the utilisation of Internet as a purchase channel: only a confident web purchaser has the awareness of the enormous amount of high-quality information he/she can collect.

According to the differences in customer's perceived risk and purchasing experience, on the one hand multi-channels allow customers to select their preferred channel, on the other hand businesses may reach customers in several ways [32].

3 Data

Eurobarometer consists in a series of surveys regularly performed on behalf of the European Commission (http://ec.europa.eu/public_opinion/index_en.htm). The aim of these surveys is to monitor the evolution of public opinion throughout the EU Member States, over a wide range of topical issues relating to the European Union. Eurobarometer results can provide well-timed and up-to-date information on the public opinion in the Member States, in order to help EU policy-maker decisions. Among the main topics investigated by means of these surveys, we can mention the EU enlargement, Euro currency, social situation, health, culture, information technology, environment and so on.

Data analysed in this paper were collected in February and March 2008 through the Eurobarometer 69.1 (ZA4743), where the major areas of interest were discrimination, radioactive waste and purchasing in the European Union. For the latter focus, respondents were asked about whether they had purchased or tried to purchase goods/services in the last 12 months, in their own and/or in other EU countries, and which purchase channels they used. Demographic, socio-economic and other background information were also collected.

In that wave, 26,746 citizens of the 27 EU member countries were interviewed. Respondents are mainly females (56 %) and 48 years old on average (ranging from 15 to 98 years of age), even if there is a large cross-country heterogeneity: in Germany, Finland, The Netherlands and Hungary the average age is higher than 50 years, while in Ireland, Italy, Latvia this value is even lower than 45 years.

About 8 % of the respondents are still studying. Among the remaining, almost half of them spent between 10 and 13 years in full-time education. Nordic countries show the largest educational levels. Half of the sample is composed by workers and about 28 % are retired. The remaining are divided between homemakers, students and unemployed. The largest proportion of employed people (more than 60 %) are in Austria, Slovakia and Sweden, the lowest in Malta, Hungary and Slovenia (less than 40 %). Baltic countries show the largest fractions of students.

Respondents are about equally divided across rural areas, small or medium towns and large towns. In Belgium, Luxembourg, Austria and Malta rural areas or villages are reported by more than half of the individuals, while the largest proportion of large town citizens is in Greece. Households are mainly composed by two individuals.

About 60 % of respondents own a PC: Scandinavian countries plus The Netherlands and Luxembourg report the highest rates of PC ownership, while the opposite applies in Bulgaria, Greece, Hungary and Romania. Similar relationships appear looking at the Internet connections, even though the overall proportion of respondents with an Internet access is about 51 %.

There is evidence of different utilisation of the purchase channels across countries and age groups. Some channels are evaluated as risky (for instance, because of frauds) by most consumers.

From Table 1, we can see that the web channel shows the largest percentages in countries with a well-known widespread presence of PC and Internet connection. Purchasing via mail is largely adopted in Great Britain, while it is marginally used in Bulgaria and Cyprus. The phone is particularly important for buying in many Central Europe countries (Germany, France, Austria, etc.), while in East Europe and Finland there is a large use of sales representatives. Apart from Luxembourg, the option of purchasing outside the national boundaries via phone or mail is not common. A very large cross-country heterogeneity is evident for the web channel (outside national boundaries).

Table 1 Distribution of the purchasing channels utilisation, across countries (%)

Country	National via internet	National via mail	National via phone	National via repres.	Abroad via internet	Abroad via mail	Abroad via phone
Belgium	22.1	11.0	26.8	7.9	14.0	2.5	3.0
Denmark	54.8	18.0	19.9	7.9	27.5	1.5	3.0
Germany	34.1	18.4	44.8	7.7	6.2	1.1	1.4
Greece	5.4	5.7	5.3	5.4	5.7	0.7	1.3
Spain	14.0	7.2	10.2	7.2	9.4	3.2	2.8
Finland	33.2	30.0	38.0	12.7	12.8	1.9	3.4
France	39.5	16.9	44.2	8.7	9.7	0.3	1.5
Ireland	19.1	11.3	11.0	8.4	18.6	2.6	3.6
Italy	12.2	8.5	10.4	8.8	5.8	3.7	4.1
Luxembourg	9.6	6.6	6.6	7.4	35.7	7.0	18.3
The Netherlands	57.8	21.2	33.2	7.5	18.5	1.6	2.3
Austria	23.5	13.5	40.2	7.3	18.2	3.7	9.4
Portugal	6.6	3.4	9.8	5.9	2.3	1.0	0.8
Sweden	59.2	27.2	33.4	10.1	20.9	0.9	2.0
Great Britain	47.1	33.8	40.0	8.6	13.8	1.4	1.7
North Ireland	39.5	18.3	32.2	6.6	10.3	1.3	1.3
Cyprus	3.0	1.2	0.8	4.9	10.1	0.2	1.2
Czech Rep.	32.0	9.7	32.8	12.6	3.5	1.0	2.3
Estonia	18.7	13.4	32.8	5.7	7.2	1.7	7.3
Hungary	9.2	7.5	17.0	12.9	0.9	0.4	0.6
Latvia	21.7	14.0	28.1	10.7	5.9	2.0	8.8
Lithuania	6.6	5.2	10.6	7.7	3.4	0.5	0.9
Malta	2.2	15.4	6.2	8.8	18.8	0.8	5.6
Poland	22.5	5.7	15.3	5.5	2.4	0.3	0.5
Slovakia	11.9	6.3	25.4	9.1	2.5	0.7	2.3
Slovenia	15.1	17.6	33.1	11.2	6.5	1.9	3.5
Bulgaria	3.7	2.5	4.8	7.5	1.0	0.4	0.5
Romania	5.6	4.8	8.5	13.3	1.1	0.8	1.3

4 The model

The approach adopted in this paper is the multilevel Latent Class (LC) analysis [37]. LC modelling is a powerful way to classify units in latent subgroups or market segments. Originally introduced by [19], LC analysis was extended by [15,20], but formalised and further developed only recently [25,39,40]. While a typical application of a LC solution is to cluster units in groups (latent classes) with similar characteristics, the multilevel strategy allows us to take into account the dependencies between the lower-level units resulting from the hierarchical data structure and to simultaneously obtain first- and second-level latent classes. In the data, second-level units are the countries and first-level units are cross-national consumers.

Let $i = 1, \ldots, I$ be the cross-national consumers, living in one of the 27 European countries, denoted by $j = 1, \ldots, 27$; in each country there are n_j consumers. Let $p =$

$1, \ldots, P$ be the investigated purchasing channel (e.g., via web within the national boundaries, via mail within the national boundaries, etc.). Thereby, $Y_{ijp} = 1$ if individual i of country j purchased a good or service through the channel p in the last 12 months, 0 otherwise. Let $\mathbf{Y_{ij}}$ and $\mathbf{Y_j}$ be the full vector of responses of individual i in country j and the full set of responses of country j, respectively. Let L be the number of lower-level LCs, H be the number of higher-level LCs. Consumer and country segment membership is represented by the discrete latent variables U_{ij} and U_j^g, respectively.

At the country level, the model specifies the marginal probability of the full response vector of country j as:

$$P(\mathbf{Y_j} = \mathbf{y_j}) = \sum_{u^g=1}^{H} P(U_j^g = u^g) \prod_{i=1}^{n_j} P(\mathbf{Y_{ij}} = \mathbf{y_{ij}} | U_j^g = u^g)$$

where $P(U_j^g = u^g)$ is the probability that group j belongs to LC u^g and $P(\mathbf{Y_{ij}} = \mathbf{y_{ij}} | U_j^g = u^g)$ is the conditional probability for the full response vector of individual i in group j conditional on the membership of group j to LC u^g. Note that the observations of the n_j respondents in each country j are assumed to be independent of one another given the country LC membership.

At the individual level, the model specifies the conditional probabilities of the full vector of responses of individual i in country j as:

$$P(\mathbf{Y_{ij}} = \mathbf{y_{ij}} | U_j^g = u^g) = \sum_{u=1}^{L} P(U_{ij} = u | U_j^g = u^g) \prod_{p=1}^{P} P(Y_{ijp} = y_{ijp} | U_{ij} = u)$$

where $P(U_{ij} = u | U_j^g = u^g)$ is the probability that individual i of group j belongs to LC u given that the group belongs to LC u^g, and $P(Y_{ijp} = y_{ijp} | U_{ij} = u)$ is the conditional density for response variable p of individual i in group j given the membership to individual-level class u. The purchases through different channels are assumed to be independent, conditional on consumer latent class membership.

Combining the two equations we obtain the following formulation, such as the one proposed by [37]:

$$P(\mathbf{Y_j} = \mathbf{y_j}) = \sum_{u^g=1}^{H} \left[P(U_j^g = u^g) \prod_{i=1}^{n_j} \left[\sum_{u=1}^{L} P(U_{ij} = u | U_j^g = u^g) \cdot \right.\right.$$
$$\left.\left. \cdot \prod_{p=1}^{P} P(Y_{ijp} = y_{ijp} | U_{ij} = u) \right]\right] \quad (1)$$

The three components of Eq. (1) are modelled as logit equations. The first component is the probability that country j belongs to a particular country segment. The second component is the probability that individual i belongs to a specific respondent segment, conditional on country segment membership: this component captures key differences between country segments. The last component is the conditional probability that individual i purchases via the channel p given his/her segment membership probabilities and it captures the main differences between respondent segments.

In a concomitant variable perspective [9,10], an important extension of the model is the possibility to include covariates to predict class membership, both at the lower and higher-level. In this application, a set of level 1 exogenous variables, X_{tij} $(t = 1, \ldots, T)$, is introduced to predict to which cluster an individual belongs to, while we do not use level 2 exogenous variables. The concomitant variables we consider in our specification are respondent characteristics such as age, gender, education, occupation, household size and area of residence. In particular, we assume that the influence of Xs on Ys is completely mediated by the latent variable U, obtaining, for $u' = 2, \ldots, L$, the following formulation:

$$P(U_{ij} = u' | U_j^g = u^g, X_{tij})) = \frac{\exp(\gamma_{0u'u^g} + \sum_{t=1}^{T} \gamma_{tu'} X_{tij})}{\sum_{u=1}^{L} \exp(\gamma_{0uu^g} + \sum_{t=1}^{T} \gamma_{tu} X_{tij})}.$$

The parameters of the multilevel LC model in Eq. (1) can be estimated by maximum likelihood (ML), obtained by an adapted EM algorithm [37]. The multilevel LC model cannot be estimated with standard software for LC analysis, we therefore use the Latent GOLD software in which the algorithm has been implemented [42]. Alternative software that can be used to estimate multilevel LC models are the R package MultiLCIRT [2], where purchasing channels can be treated as binary items, MPlus [26] and the LCA Stata plugin [18].

While the model selection issue and in particular the decision about the number of clusters is one of the main research topics in LC analysis, there has been little attention in the multilevel context (see, for example, [23,24]). The most popular set of model selection tools in LC cluster analysis are the information criteria (the lower the better), such as Akaike information criterion—AIC [1], Bayesian information criterion—BIC [31] and consistent AIC—CAIC [6]. Furthermore, criteria based on the uncertainty of classification like the estimated total number of misclassifications and entropy are also used.

Alternative model specifications have been estimated for different values of L and H. The most appropriate model has been chosen comparing the models fit, using the minimum CAIC rule as in [5]. To account for sub-optimal solutions, each model has been estimated several times with different random starting values. We retained the best solution for each specification.

5 Empirical results

The model defined in Eq. (1), including the effects of some demographic characteristics by means of concomitant variables, is estimated for different number of clusters at each level. Table 2 summarises the CAIC values obtained with these specifications. Fit of models with lower numbers of lower-level classes is not reported because they systematically show significant higher values.

The minimum CAIC rule suggests to choose the specification with nine consumer-classes and eight country-classes. However, for content reasons, we decided to compare in more

Table 2 Values of the CAIC indicator for alternative numbers of country and consumer segments

Number of consumer segments (L)	Number of country segments (H)							
	1	2	3	4	5	6	7	8
5	108,573	104,604	103,492	102,904	102,777	102,689	102,331	102,297
6	108,378	104,301	103,256	102,597	102,513	102,474	102,002	101,946
7	108,375	104,243	103,140	102,533	102,463	102,370	102,083	101,918
8	108,411	104,266	103,147	102,808	102,445	102,463	101,868	101,967
9	108,499	104,350	103,118	102,569	102,192	102,159	101,848	101,817

Table 3 Comparison of the fit of two estimated multilevel LCs models

L	H	CAIC	$R^2_{entropy,high}$	$R^2_{entropy,low}$	Classification error	Number of parameters
9	8	101,817.3	1.00	0.70	0.21	222
9	7	101,848.1	0.99	0.64	0.21	213
Δ		0.03 %	1 %	9 %	0 %	9

Table 4 Model results: consumer segments

	Consumer segments								
	1	2	3	4	5	6	7	8	9
Sizes $P(U_{ij} = u)$	0.45	0.12	0.12	0.10	0.06	0.06	0.04	0.03	0.02
Purchasing channels: Purchasing probabilities $\{P(Y_{ijp} = 1 \vert U_{ij} = u)\}$									
National via internet	0.00	0.86	0.03	0.23	0.70	0.06	0.89	0.07	0.79
National via mail	0.00	0.42	0.76	0.06	0.67	0.38	0.15	0.08	0.75
National via phone	0.01	0.07	0.25	0.05	0.85	0.10	0.22	0.11	0.56
Nnational via repres.	0.02	0.07	0.10	0.04	0.28	0.61	0.07	0.14	0.36
Abroad via internet	0.00	0.10	0.00	0.29	0.08	0.00	0.52	0.53	0.75
Abroad via mail	0.00	0.02	0.02	0.01	0.00	0.00	0.01	0.56	0.40
Abroad via phone	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.32	0.26

details both the model preferred by CAIC criteria and the model with nine consumer segments and seven country segments. Results of model fit statistics are reported in Table 3. It is true that the model with nine consumer-classes and eight country-classes could be preferred also with respect to the entropy-based R-squared measures (the nearest the value to 1 the better the model fit, see [41] for details), however adding nine parameters in the model specification leads to a modest reduction in the CAIC value (0.03 %) and the classification error does not differ between the two models. In other words, the model fit slightly increases by adding one (entire) higher-level class. Additionally, the indicators of purchasing from sales representatives and via mail or phone outside the national boundaries are better explained in the more parsimonious model. Therefore, in order to have more explicable results, we prefer to choose the more parsimonious model which better captures the less frequent purchasing channels, even if we do not respect the suggestions coming from the CAIC and entropy-based R-squared values.

Table 4, 5 and 6 present the results of the multilevel LC model with respect to the consumer segments. Table 4 shows sizes of consumer segments ($P(U_{ij} = u)$) and purchasing probabilities conditional on consumer segment ($P(Y_{ijp} = 1|U_{ij} = u)$). Table 5 reports the distribution of demographic variables used in the model (averaged across all categories of the other variables), conditional to consumer membership segment; Table 6 shows the conditional probabilities of belonging to a given consumer level segment for some specific individual profiles ($P(U_{ij} = u|X_{tij})$).

Table 4 shows that the first lower-level class is characterised by low probability for all purchase channels; therefore, we call it the *no-purchasers* class. It is a very large cluster, since its size represents half of the whole sample. This class is mainly composed by retired individuals or housewives. Even though the average age is around 52 years, about two thirds of the overall 65+ respondents belong to this cluster. This class assembles almost all (90 %)

Table 5 Conditional distribution of demographic variables, given consumer membership segment, $P(X_{tij}|U_{ij}=u)$

	Consumer segments								
	1	2	3	4	5	6	7	8	9
Gender									
Male	0.45	0.42	0.31	0.60	0.43	0.24	0.73	0.41	0.50
Female	0.55	0.58	0.69	0.40	0.57	0.76	0.27	0.59	0.50
Age									
15–24	0.08	0.21	0.05	0.26	0.09	0.17	0.15	0.15	0.12
25–34	0.10	0.22	0.08	0.24	0.17	0.20	0.28	0.17	0.20
35–44	0.14	0.23	0.12	0.21	0.23	0.23	0.27	0.24	0.30
45–54	0.17	0.17	0.16	0.14	0.22	0.20	0.17	0.18	0.20
55–64	0.19	0.11	0.22	0.09	0.17	0.12	0.10	0.15	0.12
65+	0.32	0.06	0.36	0.07	0.10	0.09	0.03	0.11	0.05
Education (years)									
5	0.05	0.00	0.01	0.00	0.00	0.02	0.00	0.01	0.00
6–9	0.29	0.07	0.26	0.04	0.09	0.15	0.01	0.11	0.04
10–13	0.48	0.49	0.51	0.43	0.47	0.57	0.18	0.51	0.44
14+	0.17	0.44	0.23	0.53	0.44	0.26	0.81	0.36	0.52
Economic status									
Retired/disabled	0.39	0.10	0.47	0.11	0.15	0.15	0.05	0.16	0.08
Employee	0.34	0.58	0.34	0.50	0.56	0.48	0.69	0.53	0.50
Self-employed	0.05	0.07	0.03	0.10	0.13	0.10	0.12	0.07	0.22
Homemaker	0.11	0.06	0.07	0.06	0.05	0.10	0.00	0.11	0.07
Student	0.05	0.15	0.03	0.20	0.05	0.11	0.12	0.10	0.09
Unemployed	0.06	0.05	0.06	0.03	0.05	0.06	0.02	0.03	0.04
Household size									
1	0.24	0.13	0.27	0.16	0.13	0.13	0.25	0.15	0.11
2	0.36	0.27	0.41	0.25	0.32	0.26	0.31	0.28	0.28
3	0.17	0.22	0.15	0.23	0.21	0.23	0.17	0.23	0.22
4+	0.23	0.38	0.18	0.36	0.34	0.38	0.27	0.34	0.38
Type of community									
Rural area or village	0.38	0.33	0.42	0.28	0.37	0.40	0.14	0.38	0.42
Small or medium city	0.36	0.40	0.38	0.37	0.43	0.35	0.38	0.39	0.37
Large city	0.26	0.27	0.20	0.35	0.19	0.25	0.48	0.22	0.21

of low educated respondents (up to 5 years of full-time education). Respondents in this class mostly live in rural areas.

In the second class, there is a high probability of using the web for national purchases, in particular among students and workers. Hence, this is the *national web purchasers* class. This cluster is characterised by a low average age (around 39 years old) and high education (nearly half of the individuals in this class have studied for more than 13 years). These purchasers mainly live in small/medium towns and in large-size households.

Table 6 Conditional distribution of consumer segments, given demographic variables, $P(U_{ij} = u|X_{tij})$

	$P(X_{tij})$	Consumer segments								
		1	2	3	4	5	6	7	8	9
Gender										
Male	0.44	0.46	0.11	0.08	0.14	0.06	0.03	0.07	0.03	0.02
Female	0.56	0.45	0.12	0.14	0.07	0.06	0.08	0.02	0.03	0.02
Age										
15–24	0.12	0.3	0.2	0.05	0.21	0.05	0.08	0.06	0.04	0.02
25–34	0.15	0.3	0.18	0.06	0.16	0.07	0.07	0.09	0.04	0.02
35–44	0.18	0.36	0.15	0.08	0.12	0.08	0.07	0.07	0.04	0.03
45–54	0.17	0.45	0.12	0.11	0.08	0.08	0.06	0.04	0.03	0.02
55–64	0.16	0.53	0.08	0.16	0.06	0.07	0.04	0.03	0.03	0.01
65+	0.22	0.67	0.03	0.19	0.03	0.03	0.02	0.01	0.02	0
Education (years)										
5	0.02	0.91	0	0.03	0	0.01	0.04	0	0.01	0
6–9	0.19	0.68	0.04	0.15	0.02	0.03	0.04	0	0.02	0
10–13	0.47	0.47	0.12	0.12	0.09	0.06	0.07	0.02	0.03	0.02
14+	0.31	0.26	0.17	0.09	0.17	0.09	0.05	0.12	0.04	0.03
Economic status										
Retired/disabled	0.28	0.63	0.04	0.19	0.04	0.03	0.03	0.01	0.02	0
Employee	0.43	0.36	0.16	0.09	0.12	0.08	0.06	0.07	0.04	0.02
Self-employed	0.07	0.34	0.11	0.05	0.15	0.12	0.08	0.08	0.03	0.06
Homemaker	0.09	0.58	0.08	0.1	0.08	0.04	0.06	0	0.04	0.01
Student	0.08	0.28	0.21	0.04	0.24	0.04	0.07	0.06	0.04	0.02
Unemployed	0.05	0.53	0.11	0.13	0.05	0.06	0.06	0.02	0.02	0.01
Household size										
1	0.2	0.53	0.08	0.15	0.08	0.04	0.04	0.06	0.02	0.01
2	0.33	0.49	0.1	0.14	0.08	0.06	0.04	0.04	0.03	0.02
3	0.19	0.42	0.13	0.09	0.12	0.07	0.07	0.04	0.04	0.02
4+	0.28	0.38	0.16	0.08	0.13	0.08	0.08	0.04	0.04	0.02
Type of community										
Rural	0.36	0.48	0.11	0.14	0.08	0.07	0.06	0.02	0.03	0.02
Small	0.38	0.44	0.13	0.12	0.1	0.07	0.05	0.05	0.03	0.02
Large	0.27	0.44	0.12	0.09	0.13	0.05	0.05	0.08	0.03	0.01

An attractive segment from a marketing point of view is the third one, the so-called *national mail purchasers* class: it has a high probability of having middle-age women who buy by mail. The other main characteristics of the individuals belonging to this class are very similar to those of the *no-purchasers*.

Consumers of the fourth class are called *potential purchasers*. On the one hand, similar to the *no-purchasers* class, this segment has quite low probabilities for all purchase channels (even if not so low as the first LC). On the other hand, the effects of demographic variables are similar to the *national web purchasers* class. In other words, they are mainly young, high educated and living in large-size households.

Classes 2, 3 and 4 have a similar size, from 10 to 12 % of the overall sample. Moreover, about 80 % of the respondents are classified in one of the first four classes.

The *national purchasers* class describes the fifth segment of consumers, who have high probability for national buying with all channels (even if the purchasing by phone is particularly important). Individuals in this class are mainly workers, well-educated and living in large-size households.

The sixth class (*in person purchasers*) is more likely to be composed by individuals who buy from sales representatives. It is worth noting the ability of the multilevel LC approach to isolate a class composed largely by consumers who exploit this specific purchasing channel. Individuals belonging to this class are more likely to be young women, medium educated and living in rural areas.

The last three classes show the smallest sizes (less than 10 % altogether). Respondents who purchase via Internet, both within and outside the national boundaries (*web purchasers*) have high probability of being included in the seventh class. They are mainly young and well-educated men: the average age is lower than 40 years, while almost 80 % have studied for more than 13 years. They live in large town, in medium-size households. The last two classes have high penetration rates for purchasing only outside the national boundaries (*international purchasers*) and purchase via all the aforementioned channels (*general purchasers*), respectively. Age and education have the strongest effect across all classes. Indeed, the respondent characteristics are quite similar across these classes, even if the former shows a bit larger presence of workers.

From a marketing point of view, classes 1 (the *no-purchasers*) and 4 (the *potential purchasers*) are the most interesting. On the one hand, both identify individuals who do not make use of any of these channels. On the other hand, their main characteristics are quite different (old and inactive people in the first class, the opposite in the second class): we can suppose a different reaction in the two groups to specific marketing stimuli. In particular, by deeply analysing the clusters we notice that PC and Internet connection ownership is widespread among the *potential purchasers*; therefore, e-commerce enterprises could plan for them fruitful strategies. Moreover, the size of these potential consumers is not so small (10 % of the overall sample). The *no-purchasers* class collects individuals who probably purchase through a traditional retail dealer and it is hard to suppose a change in such a behaviour.

Another interesting result is the definition of two classes of web purchasers, that is buying through Internet within the national boundaries only (the second class), or within and outside the national boundaries (the seventh class). This subtle distinction could lead to identify for instance consumers with higher or lower risk perception for purchasing in foreign websites, maybe related to the knowledge (or not) of foreign international languages.

As far as the model results on the higher level of the analysis are concerned, Table 7 describes the consumer segment probabilities conditional on country segments ($P(U_{ij} = u|U_j^g = u^g)$), Table 8 shows the higher-level cluster size ($P(U^g = u^g)$) and the result of the classification of the countries based on the empirical Bayesian posterior distribution [33], which graphical representation is given in Fig. 1.

Mediterranean countries, Romania and Bulgaria belong to the first segment (the largest), that is mainly composed by *no-purchasers*, even if there is a non-trivial presence of individuals who buy from sales representatives (this channel is particularly important in Romania). These countries are characterised by low degrees of diffusion of PCs and Internet connections.

The second class includes five countries (Belgium, Slovenia, Czech Republic and two Baltic countries) and the common feature is a large use of mail purchases. However, we can still observe a strong component of *no-purchasers*. It is worth noting that Baltic countries do

Table 7 Model results: conditional consumer segment probabilities ($P(U_{ij} = u|U_j^g = u^g)$)

Consumer segments	Country segments						
	1	2	3	4	5	6	7
1, no-purchasers	0.73	0.42	0.24	0.43	0.21	0.60	0.25
2, national web purchasers	0.01	0.14	0.24	0.00	0.28	0.13	0.07
3, national mail purchasers	0.01	0.21	0.23	0.02	0.11	0.09	0.29
4, potential purchasers	0.08	0.07	0.04	0.30	0.08	0.01	0.16
5, national purchasers	0.02	0.05	0.16	0.03	0.11	0.02	0.02
6, in person purchasers	0.11	0.04	0.00	0.05	0.00	0.12	0.00
7, web purchasers	0.00	0.02	0.07	0.05	0.18	0.01	0.06
8, international purchasers	0.02	0.04	0.01	0.10	0.01	0.01	0.05
9, general purchasers	0.00	0.02	0.02	0.02	0.03	0.00	0.10

Table 8 Model results: size of the higher-level LCs and country classification based on the empirical Bayesian posterior distribution

Country segments	$P(U^g = u^g)$	Countries
1	0.25	Greece, Italy, Portugal, Cyprus, Lithuania, Bulgaria, Romania
2	0.18	Belgium, Czech Republic, Estonia, Latvia, Slovenia
3	0.17	Germany, France, Finland, Great Britain, Northern Ireland
4	0.14	Spain, Ireland, Luxembourg, Malta
5	0.12	Denmark, The Netherlands, Sweden
6	0.11	Hungary, Poland, Slovakia
7	0.04	Austria

no belong to the same class, but this is not surprising thinking at the different growths and economic progress shown by these countries after their EU entry.

The third class is made up of countries with the largest probabilities of lower-level classes involving any national purchasing channel (by web, mail or phone). This class is composed by the three richest European countries, Germany, France and UK, plus Northern Ireland and Finland.

At a first sight, the fourth segment seems the less homogeneous class, because it is not easy to understand the presence of a Mediterranean country (Spain), together with the smallest European countries (Luxembourg and Malta) and Ireland. However, we can note that in this class it is more likely to find potential and international purchasers. This could be due to geographic and/or country-dimension reasons, that facilitate the relationships between these and other European and non-European countries.

Denmark, The Netherlands and Sweden belong to the fifth segment, which is characterised by the highest use of the web as purchasing channel (in terms of both *national web purchasers* and *web purchasers*). This finding is not surprising given the high rates of PC and Internet connection in the whole population of each country, as well as a high degree of knowledge of international languages, like English.

The sixth segment shows an interesting geographic proximity of the countries (Poland, Slovakia and Hungary), where the most common channels are national web and sales repre-

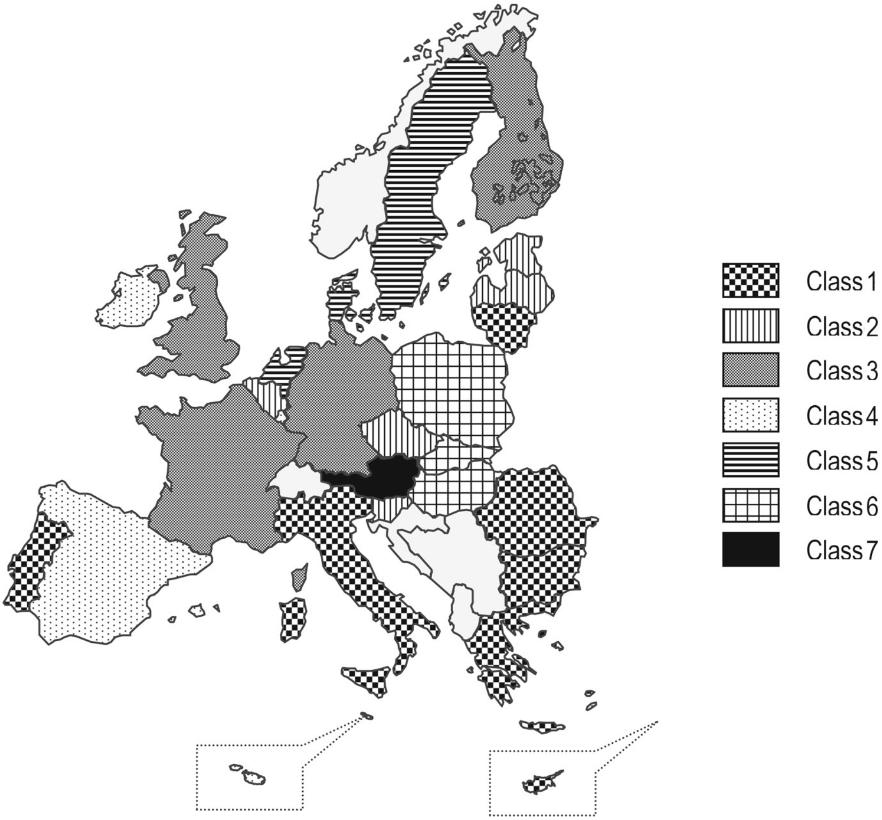

Fig. 1 Country classification based on the empirical Bayesian posterior distribution

sentatives. However, at the same time, there is still a large presence of *no-purchasers*. This class seems the most homogeneous under several points of view (e.g. socio-demographic characteristics of the individual belonging to each countries, PC and Internet diffusion, etc.).

Austria belongs to the last (country-specific) segment, which shows an interesting mix of *no-purchasers* and *national mail purchasers*, as well as the largest probability (among all seven country segments) of *general purchasers*.

6 Conclusions

[12] claims that "Euromarketing might be the name of the game for the future". Indeed, the European Union is a distinguishing example of controversial unique market. European countries are clearly separated each other according to national, cultural and economic dimensions. Recognising substantial differences across European countries might support the use of multi-domestic or multi-regional policies. However, in the last years many political and monetary efforts (i.e. the introduction of the Euro currency and the role of the European Central Bank) have been made to create a potential large single market.

In this paper, we deal with European consumers facing the purchasing process by means of difference purchasing channels. Analysing data collected in 2008 by the Eurobarometer

69.1 survey, we shed more light on consumers behaviour regarding purchasing preferences. To this aim, a multilevel latent class analysis is chosen as segmentation tool.

The overall finding is that it is difficult to treat the different European states as a unique market and most countries can be grouped in classes that follow a geographical division. Northern Europe citizens favour web purchases, probably due to the high Internet penetration rates in these countries, while consumers from Central Europe countries make use of several purchasing channels. On the other hand, in Mediterranean countries the alternatives to the traditional retail dealer are poorly exploited and this result could be due to both cultural characteristics and habits and to structural deficits in Information Technology systems (i.e. in high-speed Internet connection). European Union has to make a great effort to guarantee similar service levels across all European citizens.

At the same time, European consumers can be divided in classes having homogeneous purchasing patterns. Our analysis allows to identify some interesting consumer profiles. On the one hand, the web is the only purchasing channel for a particular group of individuals. On the other hand, another class joins consumers who need a direct contact with the product to buy. Yet, some other consumers can be grouped according to their preferences for purchasing within (but not outside) national boundaries. In these groups it is noticeable the role played by individual risk perception (i.e. to fraud), background and skills (i.e. the knowledge of international languages) or e-technology ownership in choosing one channel instead of another. The identification of such segments of consumers and their main characteristics (i.e. socio-demographic profiles) are particularly helpful for companies, in order to implement profitable marketing strategies within or across countries.

The future of this research is to investigate to what extent these classes are stable over time. Unfortunately, the lack of longitudinal data does not allow to study an updated figure of these dynamics: is the growing diffusion of Internet able to lead the web as the most important purchasing channel in alternative to the traditional retail? Could the ageing process modify the way of buying products?

In the end, it is worth noting that a potential limitation of our analysis could be related to the instrument adopted to collect the data. We could find heterogeneity in the translation process of the country-specific questionnaires or in the cross-national interpretation or understanding of the same survey questions. However, the long history and experience of the Eurobarometer surveys should prevent from or, at least, control for these potential problems.

Acknowledgements This research was partially supported by the University of Padua (Italy) with grant CPDA121180 "Statistical and econometric approach to marketing: applications and developments to customer satisfaction and market segmentation".

References

1. Akaike, H.: A new look at the statistical model identification. IEEE Trans. Autom. Control **19**, 716–723 (1974)
2. Bartolucci, F., Bacci, S., Gnaldi, M.: MultiLCIRT: an R package for multidimensional latent class item response models. Comput. Stat. Data Anal. **71**, 971–985 (2014)
3. Bassi, F.: Latent class models for marketing strategies: an application to the Italian Pharmaceutical Market. Methodol. Eur. J. Res. Methods Behav. Soc. Sci. **5**, 40–45 (2009)
4. Bhatnagar, A., Ghose, S.: A latent class segmentation analysis of e-shoppers. J. Bus. Res. **57**, 758–767 (2004)
5. Bijmolt, T.H.A., Paas, L.J., Vermunt, J.K.: Country and consumer segmentation: multi-level latent class analysis of financial product ownership. Int. J. Res. Mark. **21**, 323–340 (2004)

6. Bozdogan, H.: Model selection and Akaike's information criterion(AIC): the general theory and its analytical extensions. Psychometrika **52**, 345–370 (1987)
7. Butanay, G., Wortzel, L.H.: Distributor power versus manufacturer power: the customer role. J. Mark. **52**, 52–63 (1988)
8. Collesei, U., Casarin, F., Vescovi, T.: Internet e i cambiamenti nei comportamenti d'acquisto del consumatore. Micro Macro Mark. **1**, 33–50 (2001)
9. Dayton, C.M., Macready, G.B.: Concomitant-variable latent-class models. J. Am. Stat. Assoc. **83**, 173–178 (1988)
10. Formann, A.K.: Mixture analysis of multivariate categorical data with covariates and missing entries. Comput. Stat. Data Anal. **51**, 5236–5246 (2007)
11. Friedman, L.G., Furey, T.R.: The Channel Advantage. Butterworth-Heinemann, Burlington (2003)
12. Ganesh, J.: Converging trends within the European Union: insights from an analysis of diffusion patterns. J. Int. Mark. **6**, 32–48 (1988)
13. Gnaldi, M., Bacci, S., Bartolucci, F.: A multilevel finite mixture item response model to cluster examinees and schools. Adv. Data Anal. Classif. (2015). doi:10.1007/s11634-014-0196-0
14. Henry, K.L., Muthén, B.: Multilevel latent class analysis: an application of adolescent smoking typologies with individual and contextual predictors. Struct. Equ. Model. **17**, 193–215 (2010)
15. Goodman, L.A.: Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika **61**, 215–231 (1974)
16. Keng, K., Tang, Y., Ghose, S.: Typology of online shoppers. J. Consum. Mark. **20**, 139–156 (2003)
17. Kumar, V., Ganesh, J., Echambadi, R.: Cross-national diffusion research: what do we know and how certain are we? J. Prod. Innov. Manag. **15**, 255–268 (1998)
18. Lanza S.T., Dziak J.J., Huang L., Wagner A.T., Collins L.M.: LCA Stata Plugin Users' Guide (Version 1.2). University Park: The Methodology Center, Penn State (2015)
19. Lazarsfeld, P.F.: The logical and mathematical foundation of latent structure analysis and the interpretation and mathematical foundation of latent structure analysis. In: Stouffer, S.A., et al. (eds.) Measurement and Prediction, pp. 362–472. Princeton University Press, Princeton (1950)
20. Lazarsfeld, P.F., Henry, N.W.: Latent Structure Analysis. Houghton Mifflin, Boston (1968)
21. Lemmens, A., Croux, C., Dekimpe, M.G.: Consumer confidence in Europe: united in diversity? Int. J. Res. Mark. **24**, 113–127 (2007)
22. Levitt, T.: The globalization of markets. Harv. Bus. Rev. **61**, 92–102 (1983)
23. Lukočienė, O., Vermunt, J.K.: Determining the number of components in mixture models for hierarchical data. In: Fink, A., Berthold, L., Seidel, W., Ultsch, A. (eds.) Advances in Data Analysis, Data Handling and Business Intelligence, pp. 241–249. Springer, Berlin (2010)
24. Lukočienė, O., Varriale, R., Vermunt, J.K.: The simultaneous decision(s) about the number of lower- and higher-level classes in multilevel latent class analysis. Sociol. Methodol. **40**, 247–283 (2010)
25. Magidson, J., Vermunt, J.K.: Latent class models for clustering: a comparison with K-means. Can. J. Mark. Res. **20**, 36–43 (2002)
26. Muthén L.K., Muthén B.O.: Mplus Users Guide, Seventh Edition. Los Angeles, Muthén & Muthén (2012)
27. Mutz, R., Daniel, H.D.: University and student segmentation: multilevel latent-class analysis of students' attitudes towards research methods and statistics. Brit. J. Educ. Psychol. **83**, 280–304 (2013)
28. Paccagnella, O., Varriale, R.: Asset ownership of the elderly across Europe: a multilevel latent class analysis to segment countries and households. In: Torelli, N., Pesarin, F., Bar-Hen, A. (eds.) Advances in Theoretical and Applied Statistics, pp. 383–393. Springer, Berlin (2013)
29. Payne, A.F.T., Frow, P.: A strategic framework for customer relationship management. J. Mark. **69**, 167–176 (2005)
30. Saunders C.: Multiple channel buyers worth pursuing, Internet advertising report, January 28. http://www.internetnews.com (2002)
31. Schwarz, G.: Estimating the dimention of a model. Ann. Stat. **6**, 461–464 (1978)
32. Sharma, A., Mehrotra, A.: Choosing an optimal channel mix in multichannel environments. Ind. Mark. Manag. **36**, 21–28 (2006)
33. Skrondal, A., Rabe-Hesketh, S.: Generalized Latent Variable Modeling: Multilevel, Longitudinal, and Structural Equation Models. Chapman and Hall/CRC, Boca Raton (2004)
34. Solomon, M.R.: Consumer Behavior: Buying, Having and Being, 10th edn. Pearson Prentice Hall, Upper Saddle River (2012)
35. Steenkamp, J.B.E.M.: The role of national culture in international marketing research. Int. Mark. Rev. **18**, 30–44 (2001)
36. Steenkamp, J.B.E.M., Ter Hofstede, F.: International market segmentation. Int. J. Res. Mark. **19**, 185–213 (2002)
37. Vermunt, J.K.: Multilevel latent class models. Sociol. Methodol. **33**, 213–239 (2003)

38. Vermunt, J.K.: Latent class and finite mixture models for multilevel data sets. Stat. Methods Med. Res. **17**, 33–51 (2008)
39. Vermunt, J.K., Magidson, J.: Latent class cluster analysis. In: Hagenaars, J.A., McCutcheon, A.L. (eds.) Applied Latent Class Analysis, pp. 89–106. Cambridge University Press, Cambridge (2002)
40. Vermunt, J.K., Magidson, J.: Latent class models for classification. Comput. Stat. Data Anal. **41**, 531–537 (2003)
41. Vermunt, J.K., Magidson, J.: Technical Guide for Latent GOLD 4.0: Basic and Advanced, Statistical Innovations Inc. Statistical Innovations Inc., Belmont (2005)
42. Vermunt, J.K., Magidson, J.: LG-Syntax Users Guide: Manual for Latent GOLD 4.5 Syntax Module. Statistical Innovations Inc., Belmont (2008)
43. Wollace, D.W., Giese, J.L., Johnson, J.L.: Customer retailer loyalty in the context of multiple channel strategies. J. Retail. **80**, 249–263 (2004)