

# A latent class analysis of the public attitude towards the euro adoption in Poland

Ewa Genge

Received: 8 December 2012 / Revised: 2 September 2013 / Accepted: 24 October 2013 /  
Published online: 16 November 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** Latent class analysis can be viewed as a special case of model-based clustering for multivariate discrete data. It is assumed that each observation comes from one of a number of classes, groups or subpopulations, with its own probability distribution. The overall population thus follows a finite mixture model. When observed, data take the form of categorical responses—as, for example, in public opinion or consumer behavior surveys it is often of interest to identify and characterize clusters of similar objects. In the context of marketing research, one will typically interpret the latent number of mixture components as clusters or segments. In fact, LC analysis provides a powerful new tool to identify important market segments in target marketing. We used the model based clustering approach for grouping and detecting inhomogeneities of Polish opinions on the euro adoption. We analyzed data collected as part of the Polish General Social Survey using the R software.

**Keywords** Latent class analysis · Mixture model · Categorical data · Euro adoption

**Mathematics Subject Classification** 62P20 · 6207 · 91B82

## 1 Introduction

“Ten years ago, on 1 January 2002, euro banknotes and coins were introduced in 12 Member States of the European Union. The introduction of the euro cash was an unprecedented challenge, but it went smoothly, and billions of banknotes and coins started to circulate in a matter of days. Five more Member States have adopted the

---

E. Genge (✉)  
Department of Statistics, University of Economics in Katowice, Bogucicka 14,  
40–226 Katowice, Poland  
e-mail: ewa.genge@ue.katowice.pl

euro in recent years, so a total of 17 Member States—and 332 million people—now use the currency. It has become a symbol of Europe, and the banknotes and coins have become a part of our daily lives” (Mario Draghi, President of the ECB).

On the basis of the Eurobarometer surveys on individuals’ and firms’ perception of the currency changeover in 2002, 60 % of people considered that the changeover to the euro would bring more advantages than disadvantages. Moreover, a large majority said they felt more European thanks to the euro. Four out of five individuals felt that the changeover to the euro had gone well or very well. Lastly, over two thirds of the general public were happy that the euro was their currency and it was only in Germany, Greece and Austria that there was a higher proportion of dissatisfied individuals.

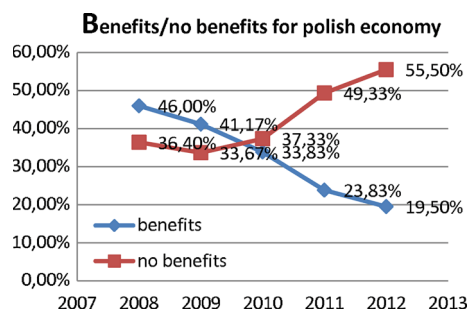
The crisis has had a major impact on EU politics, leading to power shifts in several European countries, most notably in Greece, Ireland, Italy, Portugal and Spain. There are continuous speculations on a possible breakup of the Eurozone.

Although the Eurozone is being rocked by the serious crisis, Poland declares it still wants to adopt the European Union’s common currency. Before the accession, in January 2002, Poles were very enthusiastic about the single currency. Exactly at that time, the euro was introduced in the form of banknotes and coins in the first twelve European countries. However, negative publicity surrounding the perceptions of prices in euro had not appeared then on a full scale in the media yet. According to the Flash Eurobarometer data, however, in the first years after the introduction of the physical euro the number of euro sceptics outweighed the number of euro enthusiasts in Poland.

We present Fig. 1 to see the evolution of public support for the euro introduction in years 2008–2012. The figure was prepared based on to the Polish General Social Survey (GSS) and the well-known Polish research agency TNS OBOP data.

The year 2008 was a very special time in the Polish euro history. At this time, the so-called Greek crisis began. It dramatically changed the situation on the European financial markets and triggered the process of reforming the economic governance of the EU. Before the announcement of Greece’s financial problems, the euro was rather considered as a “safe harbour” or “save haven” amid the global financial turmoil in the aftermath of the Lehman Brothers’ spectacular fall (see [Osinska and Toroj 2012](#)). We argue, however, that the outlook apparently reversed later on, when the sovereign debt problems worsened across Europe. In 2011, we faced sovereign debt crises of several euro area countries (Greece and, to a lesser extent, Ireland and Portugal).

**Fig. 1** The public attitude towards the euro and its benefits for the Polish economy in years 2008–2012



A higher share of euro supporters in 2008 could probably be linked to an association of the euro with a “safe harbour” idea at the beginning of the financial turmoil. In 2009, enthusiasm was further fostered by the successful euro changeover in Poland’s neighbour—Slovakia.

The year 2008 was also very important for Poland because on 10 September Polish Prime Minister Donald Tusk gave a speech and announced the ruling government’s objective to join the Eurozone in 2012 of the launch of an economic forum in the Polish resort of Krynica-Zdroj. He was also aware that the Polish constitution would need to be changed first and Poland would have to join the ERM 2 before the second quarter of 2009, a target date that was still very aggressive. Yet, on 28 October 2008 the Polish government confirmed their plan to join the Eurozone in January 2012. There were also many political speculations about the date of the euro adoption.

Those developments attracted extensive media coverage and at least in Poland were one of the reasons for the revision of the euro adoption timetable. As a consequence, the inevitable need for financial assistance for the peripheral euro area countries may have impacted on the public support for the euro not only in euro area countries, but also in non-EU states that were prepared to join the club in the future (nowadays we are talking about the convergence criteria to fulfill, not about the official target date). Therefore, our interest in investigating the determinants of the euro support particularly focuses on the year 2008.

The aim of this paper is to group and detect inhomogeneities of Polish opinions on the euro adoption using the model-based clustering approach. The article is organized as follows. First, we described the standard mixture model, its parameter and standard errors estimation. Subsequently, we presented the extension of the basic model which permits the inclusion of covariates to predict latent class membership. We discussed three-step and one-step approaches for LC modeling with covariates. Then we reported the problem of model selection and goodness of fit criteria. Subsequently, we presented an empirical application. The paper ends with a summary of the main results of the research and a discussion of possible directions for future research.

## 2 Mixture models

### 2.1 Mixture models

Latent class analysis was proposed by [Lazarsfeld \(1950a, b\)](#) and [Lazarsfeld and Henry \(1968\)](#) and can be viewed as a special case of model-based clustering for multivariate discrete data. In model-based clustering it is assumed that each observation comes from one of a number of classes (groups) and models each with its own probability distribution ([Wolfe 1963](#); [McLachlan and Peel 2000](#); [Banfield and Raftery 1993](#)). Because the unobserved latent variable is nominal, the latent class model is actually a type of finite mixture model.

Finite mixture models are a popular technique for modeling unobserved heterogeneity or approximating general distribution functions. They are used in a lot of different areas such as astronomy, biology, economic, marketing or medicine. An overview of mixture models is given in [Titterton et al. \(1985\)](#) or [McLachlan and Peel \(2000\)](#).

The mixture is assumed to consist of  $u$  components where each component follows a parametric distribution. Each component has a weight assigned which indicates the a-priori probability for an observation to come from this component and the mixture distribution is given by the weighted sum over the  $u$  components. The mixture model is given by

$$f(\mathbf{x}_i|\boldsymbol{\theta}) = \sum_{s=1}^u \pi_s f_s(\mathbf{x}_i, \boldsymbol{\theta}_s), \tag{1}$$

where

- $f_s$ , density function of component  $s$ ;
- $\mathbf{x}_i$ , vector of manifest variables  $\mathbf{x}_i = [\mathbf{x}_{i1}, \dots, \mathbf{x}_{im}]$ ;
- $\pi_s$ , the prior probability of component  $s$  ( $\pi_s \in (0, 1)$ ;  $\sum_s^u \pi_s = 1$ );
- $\boldsymbol{\theta}_s$ , the component specific parameter vector for the latent class  $s$ ;
- $\boldsymbol{\theta}$ , the vector of all parameters for the latent class model (mixture).

The latent class models approximate the observed joint distribution of the manifest variables as the weighted sum of finite number,  $u$ , of the cross-classification tables. The probability in each cell of the component table is simply the product of the respective class-conditional marginal probabilities. A weighted sum of these component tables forms an approximation (density estimate) of the distribution of cases across the cells of the observed table. Observations with similar sets of responses on the manifest variables ( $\mathbf{x}_i$ ) will tend to cluster within the same latent classes.

The  $s$ th component of the mixture given in formula (1) can be written as

$$f_s(\mathbf{x}_i|\boldsymbol{\theta}_s) = \prod_{j=1}^m \prod_{h=1}^{l_j} (\theta_{sjh})^{\mathbf{x}_{ijh}}, \tag{2}$$

where  $\mathbf{x}_i = (x_{ijh}; j = 1, \dots, m; h = 1, \dots, l_j; i = 1, \dots, n)$ <sup>1</sup>,  $\boldsymbol{\theta}_s = (\theta_{ijh}; j = 1, \dots, m; h = 1, \dots, l_j; i = 1, \dots, n)$  and <sup>2</sup> formula is a product of conditionally independent multinomial distributions of parameters  $\boldsymbol{\theta}_{sj}$ .

The probability density function across all classes can be also written as the weighted sum:

$$P(\mathbf{x}_i|\boldsymbol{\theta}) = \sum_{s=1}^u \pi_s \prod_{j=1}^m \prod_{h=1}^{l_j} (\theta_{sjh})^{\mathbf{x}_{ijh}}. \tag{3}$$

The parameters estimated by the latent class model are  $\pi_s$  and  $\boldsymbol{\theta}_{sjh}$ .

Given estimates  $\hat{\pi}_s$  and  $\hat{\theta}_{sjh}$  of  $\pi_s$  and  $\theta_{sjh}$ , the posterior probability (that each individual belongs to each class, conditional on the observed values of the manifest variables), can be calculated using Bayes' formula:

<sup>1</sup>  $x_{ijh} = 1$  if the object  $i$  belongs to the category  $h$  of the variable  $j$ . The total number of categories is given by  $l = \sum_{j=1}^m l_j$ , then the data is defined by the  $n$  by  $m$  matrix.

$$\hat{P}(s|\mathbf{x}_i) = \frac{\pi_s f(\mathbf{x}_i; \hat{\theta}_s)}{\sum_{q=1}^u \pi_q f(\mathbf{x}_i; \hat{\theta}_q)}. \tag{4}$$

The number of independent parameters estimated by the latent class model increases rapidly with  $u, j, l_j$ . The number of parameters is equal  $u \sum_j (l_j - 1) + (u - 1)$ . If this number exceed either the total number of observation, or one fewer than the total number of cells in the cross-classification table of the manifest variables, then the latent class model will be unidentified.

### 2.2 Parameter and standard error estimation

The parameters of the latent class model are usually estimated by maximum likelihood using the Expectation–Maximization (EM) algorithm (Dempster et al. 1977). The log-likelihood of the latent class model is given by

$$\ln L = \sum_{i=1}^n \ln \sum_{s=1}^u \pi_s \prod_{j=1}^m \prod_{h=1}^{l_j} (\theta_{sjh})^{x_{ijh}}. \tag{5}$$

Each EM iteration consists of two steps—an E step and an M step:

- E step—calculation of the “missing” class membership probabilities using Eq. 4,
- M step—maximization step, the parameter of estimates are updated by maximizing the log-likelihood function given these posterior  $\hat{P}(s|\mathbf{x}_i)$  with

$$\hat{\pi}_s^{new} = \frac{1}{n} \sum_{i=1}^n \hat{P}(s|\mathbf{x}_i). \tag{6}$$

as the new prior probabilities and

$$\hat{\theta}_{sj}^{new} = \frac{\sum_{i=1}^n \mathbf{x}_{ij} \hat{P}(s|\mathbf{x}_i)}{\sum_{i=1}^n \hat{P}(s|\mathbf{x}_i)}, \tag{7}$$

as the new class-conditional outcome probabilities.<sup>2</sup> The algorithm repeats these steps, assigning the new to old values, until the log-likelihood function reaches a maximum and ceases to increment beyond some arbitrarily small value.

Depending on the initial values chosen for  $\hat{\pi}_s$  and  $\hat{\theta}_{jsh}$ , and the complexity of the latent class model being estimated, the EM algorithm may only find a local maximum of the log-likelihood function. For this reason, in attempt to find the global maximizer, usually a latent class model is re-estimated.

Standard errors of the estimated class-conditional response probabilities and the mixing parameters are estimated most often using the empirical observed information matrix (Meilijson 1989; McLachlan and Peel 2000), which equals:

<sup>2</sup>  $\hat{\theta}_{sj}^{new}$  is the vector of length  $l_j$  of class- $s$  conditional outcome probabilities for the  $j$ th manifest variable; and  $\mathbf{x}_{ij}$  is the  $n \times l_j$  matrix of observed outcomes  $x_{ijh}$  on that variable.

$$\mathbf{I}_e(\hat{\Psi}; \mathbf{x}) = \sum_{i=1}^n \mathbf{s}(\mathbf{x}_i; \hat{\Psi}) \mathbf{s}^T(\mathbf{x}_i; \hat{\Psi}), \tag{8}$$

where  $\mathbf{s}(\mathbf{x}_i; \hat{\Psi})$  is the score function with respect to the vector of parameters  $\Psi$  for  $i$ th observation, evaluated at the maximum likelihood estimate  $\hat{\Psi}$

$$\mathbf{s}(\mathbf{x}_i; \Psi) = \sum_{s=1}^u p_{is} \partial \left\{ \ln \pi_s + \sum_{j=1}^m \sum_{h=1}^{l_j} \mathbf{x}_{ijh} \ln \theta_{sjh} \right\} / \partial \Psi, \tag{9}$$

where  $p_{is} = \hat{P}(s|\mathbf{x}_i)$  is the posterior probability that observation  $i$  belongs to class  $s$  (Eq. 4). The covariance matrix of the parameter estimates is then approximated by the inverse of  $\mathbf{I}_e(\hat{\Psi}; \mathbf{x})$ .

Because of the constraint across each manifest variables (the sum-to-one constraint), it is useful to reparameterize the score function in terms of log-ratios  $\phi_{jsh} = \ln(\theta_{jsh}/\theta_{js1})$  for given outcome variable  $j$  and class  $s$ . Then, for the  $t$ th response on the  $b$ th item in the  $q$ th class,

$$\mathbf{s}(\mathbf{x}_i; \phi_{bqt}) = p_{iq}(\mathbf{x}_{ibt} - \theta_{bqt}). \tag{10}$$

Denoting  $\omega_s = \ln(\pi_s/\pi_1)$ , then for log-ratio corresponding to the  $q$ th mixing parameter:

$$\mathbf{s}(\mathbf{x}_i; \omega_q) = p_{iq} - \pi_q. \tag{11}$$

To transform the covariance matrix of these log-ratios back to the original units of  $\boldsymbol{\pi}$  and  $\boldsymbol{\theta}$  the delta method is applied. It is assumed that  $g(\phi_{jsh}) = \theta_{jsh} = e^{\phi_{jsh}} / \sum_t e^{\phi_{jst}}$ . Denoting  $Var(\hat{\phi})$  as submatrix of the inverse of  $\mathbf{I}_e(\hat{\Psi}; \mathbf{x})$  corresponding to the  $\phi$  parameters, then

$$Var(g(\hat{\phi})) = g'(\phi) Var(\hat{\phi}) g'(\phi)^T \tag{12}$$

where  $g'(\phi)$  is the Jacobian consisting of elements

$$\frac{\partial g(\phi_{jsh})}{\partial \phi_{bqt}} = \begin{cases} 0 & \text{if } q \neq s, \\ 0 & \text{if } q = s \text{ but } b \neq j, \\ -\theta_{jsh}\theta_{jst} & \text{if } q = s \text{ and } b = j \text{ but } t \neq h, \\ \theta_{jsh}(1 - \theta_{jsh}) & \text{if } q = s \text{ and } b = j \text{ and } t = h. \end{cases} \tag{13}$$

Similarly, for the mixing parameters  $h(\omega_s) = \pi_s = e^{\omega_s} / \sum_q e^{\omega_q}$ . Taking as  $Var(\hat{\omega})$  the submatrix of the inverse of  $\mathbf{I}_e(\hat{\Psi}; \mathbf{x})$  corresponding to the  $\omega$  parameters, then

$$Var(h(\hat{\omega})) = h'(\omega) Var(\hat{\omega}) h'(\omega)^T \tag{14}$$

where  $h'(\omega)$  is the Jacobian consisting of elements

$$\frac{\partial h(\omega_s)}{\partial \omega_q} = \begin{cases} -\pi_s \pi_q & \text{if } q \neq s, \\ -\pi_s(1 - \pi_s) & \text{if } q = s. \end{cases} \quad (15)$$

Standard errors of each parameter estimate are equal to the square root of the values along the main diagonal of covariance matrices  $Var(\theta)$  and  $Var(\pi)$ .

### 3 Covariates in the latent class model

#### 3.1 Mixture models with covariates

The basic latent class model (Eq. 1) do not include the covariates ( $\mathbf{z}_i$ ). We presented the extended version of the latent class model in which the covariates permit to predict latent class membership [whereas in the basic model, every response has the same probability of belonging to each latent class (see i.e. [Clogg 1981](#); [Vermunt 1997](#); [Bandeem-Roche et al. 1997](#); [Linzer and Lewis 2011](#))].

$$f(\mathbf{x}_i, \mathbf{z}_i | \theta) = \sum_{s=1}^u \pi_s(\mathbf{z}_i, \alpha) f_s(\mathbf{x}_i, \theta_s), \quad (16)$$

where:

$f_s$ , density function of component  $s$ ;

$\mathbf{x}_i$ , vector of manifest variables  $\mathbf{x}_i = [\mathbf{x}_{i1}, \dots, \mathbf{x}_{im}]$ ;

$\mathbf{z}_i$ , vector of concomitant variables (covariates)  $\mathbf{z}_i = [\mathbf{z}_{i1}, \dots, \mathbf{z}_{if}]$ ;

$\pi_s$ , the prior probability of component  $s$  ( $\pi_s(\mathbf{z}_i, \alpha) \in (0, 1)$ ;  $\sum_s^u \pi_s(\mathbf{z}_i, \alpha) = 1$ );

$\theta_s$ , the component specific parameter vector for the latent class  $s$ ;

$\theta$ , the vector of all parameters for the latent class model (mixture).

We would like to stress that there is some confusion over the name of the latent class models with covariates. Sometimes this kind of the mixture model is called the latent class regression model, to refer to latent class models in which the probability of latent class membership is predicted by the covariates ([Linzer and Lewis 2011](#); [Vermunt and Magidson 2002](#)). There is also another approach, i.e. to use the name of latent class regression model to refer to regression models in which the dependent variable is partitioned into latent classes as part of estimating the regression model. More than one regression are simultaneously fitted to the data when the latent data partition is unknown ([Grün and Leisch 2008](#)). In this article we would rather prefer the name latent class model with covariates or concomitant-variable latent class model.

There are many techniques of including the covariates in the latent class model (i.e. see [Vermunt 2010](#)). In the empirical part of this article we applied a so-called “one-step” technique for estimating the effects of covariates, because the coefficients on the covariates are estimated simultaneously as a part of the latent class model ([Dayton and Macready 1988](#); [Hagenaars and McCutcheon 2002](#)).

It is assumed that the mixing proportions ( $\pi_{si}$ ) are free to vary, but the constraint  $\sum \pi_{si} = 1$  must be fulfilled.<sup>3</sup> In `pOLCA` package of `R` a generalized (multinomial) model logit link function for the effects of covariates on the priors is employed (Agresti 2002).

Frequently, the first latent class is a “reference” class and assumes that the log-odds of the latent class membership priors with respect to that class are linear functions of the covariates. Let  $\alpha_s$  denote the vector of coefficients corresponding to the  $s$ th latent class. With  $f$  covariates, the  $\alpha_s$  have length  $f + 1$ ; this is one coefficient on each of the covariates plus a constant. Because the first class is used as the reference,  $\alpha_s = 0$  is fixed by definition. Then,

$$\begin{aligned} \ln(\pi_{2i}/\pi_{1i}) &= \mathbf{z}_i \alpha_2^T \\ \ln(\pi_{3i}/\pi_{1i}) &= \mathbf{z}_i \alpha_3^T \\ &\vdots \\ \ln(\pi_{ui}/\pi_{1i}) &= \mathbf{z}_i \alpha_u^T \end{aligned}$$

we can get

$$\pi_{si} = \pi_s(\mathbf{z}_i; \alpha) = \frac{e^{\mathbf{z}_i \alpha_s^T}}{\sum_{q=1}^u e^{\mathbf{z}_i \alpha_q^T}}. \tag{17}$$

The parameters estimated by the latent class model with covariates are the  $u - 1$  vectors of coefficients  $\alpha_s$  and, as in the basic latent class model, the class-conditional outcome probabilities  $\theta_{sjh}$ . Given estimates  $\hat{\alpha}_s$  and  $\hat{\theta}_{sjh}$  of these parameters, the posterior class membership probabilities in the latent class model with covariates are obtained by:

$$\hat{P}(s|\mathbf{z}_i \mathbf{x}_i) = \frac{\pi_s(\mathbf{z}_i; \hat{\alpha}) f(\mathbf{x}_i; \hat{\theta}_s)}{\sum_{q=1}^u \pi_q(\mathbf{z}_i; \hat{\alpha}) f(\mathbf{x}_i; \hat{\theta}_q)}. \tag{18}$$

The number of parameters estimated by this latent class model increases rapidly with  $u, j, l_j, f$ . The number of parameters is equal  $u \sum_j (l_j - 1) + (f + 1)(u - 1)$ . If this number exceed either the total number of observation, or one fewer than the total number of cells in the cross-classification table of the manifest variables, then the latent class model will be unidentified.

### 3.2 Parameter and standard error estimation

The parameters of the latent class model with covariates are also usually estimated by maximum likelihood using the Expectation–Maximization (EM) algorithm (Dempster et al. 1977). Similarly, each EM iteration consists of two steps: an E step and an M step:

<sup>3</sup> The mixing proportions in the latent class models with covariates are denoted as  $\pi_{si}$ , to reflect the fact that these priors are now free to vary by individual, see also Eq. 17.



- E step—estimation of the posterior class probabilities for each observation and,
- M step—maximization of the log-likelihood for each component separately using the posterior probabilities as weights.

The log-likelihood of the latent class model with covariates is given by

$$\ln L = \sum_{i=1}^n \ln \sum_{s=1}^u \pi_s(\mathbf{z}_i; \boldsymbol{\alpha}) \prod_{j=1}^m \prod_{h=1}^{l_j} (\theta_{s_jh})^{\mathbf{x}_{ijh}}. \tag{19}$$

In `poLCA` package of R, in order to find the values of  $\hat{\boldsymbol{\alpha}}_s$  and  $\hat{\theta}_{s_jh}$  that maximize the function given by Eq. 19 a modified EM algorithm with a Newton-Raphson step is applied (see [Bandeem-Roche et al. 1997](#)). The estimation begins with initial values of  $\hat{\boldsymbol{\alpha}}'_s$  and  $\hat{\theta}'_{s_jh}$  to calculate posterior probabilities  $\hat{P}(s|\mathbf{z}_i \mathbf{x}_i)$ . The covariates coefficients are computed according to the formula:

$$\boldsymbol{\alpha}_s^{new} = \hat{\boldsymbol{\alpha}}'_s + (-\mathbf{D}_{\boldsymbol{\alpha}}^2 \log L)^{-1} \mathbf{D}_{\boldsymbol{\alpha}} \log L \tag{20}$$

where  $\mathbf{D}_{\boldsymbol{\alpha}}$  is the gradient and  $\mathbf{D}_{\boldsymbol{\alpha}}^2$  the Hessian matrix with respect to  $\boldsymbol{\alpha}$ . The new parameters of  $\hat{\theta}_{s_jh}$  are updated as

$$\hat{\theta}_{s_j}^{new} = \frac{\sum_{i=1}^n \mathbf{x}_{ij} \hat{P}(s|\mathbf{z}_i \mathbf{x}_i)}{\sum_{i=1}^n \hat{P}(s|\mathbf{z}_i \mathbf{x}_i)}. \tag{21}$$

These steps are repeated until convergence, assigning the new parameter estimates to the old ones in each iteration of the algorithm (see details in [Bandeem-Roche et al. 1997](#)).

When employing this estimation algorithm, different initial parameter values may lead to different local maxima of the log-likelihood function. For this reason, the algorithm should be repeated handful of times. The E and M steps are repeated until the likelihood improvement falls under a pre-specified threshold or a maximum number of iterations is reached.<sup>4</sup>

Standard errors (for latent class models with covariates) are obtained just as for models without covariates: the empirical observed information matrix is calculated. First, the score function (Eq. 9) is generalized so that

<sup>4</sup> Depending on the initial values and the complexity of the latent class model being estimated, the EM algorithm may only find a local maximum of the log-likelihood function, rather than the desired global maximum. For this reason it is always advisable to re-estimate a particular model a couple of times, in an attempt to find the global maximizer to be taken as the maximum likelihood solution. The maximum number of iterations through which the estimation algorithm will cycle in `poLCA` package of R equals 1000. A tolerance value for judging when convergence has been is  $1e - 10$ . When the one-iteration change in the estimated log-likelihood is less than that threshold, the estimation algorithm stops updating and considers the maximum log-likelihood to have been found.

$$s(\mathbf{z}_i; \mathbf{x}_i; \Psi) = \sum_{s=1}^u p_{is} \partial \left\{ \ln \pi_s(\mathbf{z}_i; \boldsymbol{\alpha}) + \sum_{j=1}^m \sum_{h=1}^{l_j} \mathbf{x}_{ijh} \ln \theta_{sjh} \right\} / \partial \Psi, \tag{22}$$

where  $p_{is}$  denote posterior probabilities. Since this function is no different than Eq. 9 in terms of  $\boldsymbol{\theta}$  parameters, the score function  $\mathbf{s}(\mathbf{z}_i; \mathbf{x}_i; \phi_{bqt}) = \mathbf{s}(\mathbf{x}_i; \phi_{bqt})$  (see Eq. 10), and the covariance matrix  $Var(\boldsymbol{\theta})$  may be calculated in the similar way as for models without covariates.

However, the priors ( $\pi_{is}$ ) are free to vary by individual as a function of some set of coefficient  $\boldsymbol{\alpha}$  (Eq. 17) letting  $q$  index classes and  $p$  index covariates:

$$s(\mathbf{z}_i; \mathbf{x}_i; \boldsymbol{\alpha}_{qp}) = \mathbf{z}_{ip}(p_{iq} - \pi_{iq}). \tag{23}$$

The standard errors of the coefficients  $\boldsymbol{\alpha}$  are equal to square root of the values along the main diagonal of the submatrix of the inverse of the empirical observed information matrix corresponding to the  $\boldsymbol{\alpha}$  parameters.

To obtain the covariance matrix of the prior parameters  $\pi_s$  which is the average value across all observations of the priors  $\pi_{is}$ , the delta method is applied (see details in Bandeen-Roche et al. 1997; Linzer and Lewis 2011, pp 5, 6, 9). If

$$h(\boldsymbol{\alpha}_s) = \pi_s = \frac{1}{n} \sum_i \left( \frac{e^{\mathbf{z}_i \boldsymbol{\alpha}_s^T}}{\sum_{q=1}^u e^{\mathbf{z}_i \boldsymbol{\alpha}_q^T}} \right) \tag{24}$$

then

$$Var(h(\hat{\boldsymbol{\alpha}})) = h'(\boldsymbol{\alpha}) Var(\hat{\boldsymbol{\alpha}}) h'(\boldsymbol{\alpha})^T \tag{25}$$

where  $h'(\boldsymbol{\alpha})$  is a Jacobian with elements

$$\frac{\partial h(\boldsymbol{\alpha}_s)}{\partial \boldsymbol{\alpha}_{qp}} = \begin{cases} 1/n \sum_i \mathbf{z}_{ip} (-\pi_{is} \pi_{iq}) & \text{if } q \neq s, \\ 1/n \sum_i \mathbf{z}_{ip} (\pi_{is} (1 - \pi_{is})) & \text{if } q = s. \end{cases} \tag{26}$$

### 4 Model selection

In order to select the optimal clustering model several measures have been proposed (see i.e. McLachlan and Peel 2000). The performance of some of these criteria were compared by Biernacki et al. (1999). In general, BIC was found to be consistent under correct specification of the component densities (Kass and Raftery 1995; Keribin 1998) and gave good results in a range of applications (i.e. Fraley and Raftery 2002; Stanford and Raftery 2000; Witek 2010, 2011b).

The most popular in the latent class analysis information criteria are available in `pOLCA` package of `R`: BIC (*Bayesian Information Criterion*) and AIC (*Akaike Information Criterion*). For this reason BIC and AIC criteria will be used in further analysis. The BIC and AIC criteria are defined as

$$BIC_S = -2 \log f(\mathbf{x}_i | \hat{\boldsymbol{\theta}}_S, M_S) + v_S \log(n), \quad (27)$$

$$AIC_S = -2 \log f(\mathbf{x}_i | \hat{\boldsymbol{\theta}}_S, M_S) + 2v_S, \quad (28)$$

where  $\log f(\mathbf{x}_i | \hat{\boldsymbol{\theta}}_S, M_S)$  is the maximized loglikelihood for the model  $M_S$ ,  $v_S$  is the number of parameters to be estimated in that model and  $n$  is the number of observations in the data. The first term of the criteria measures the goodness-of-fit, whereas the second term penalizes model complexity. One selects the model that minimizes the information criteria value.

## 5 Example

The analyses presented below are based on  $n = 648$  adolescents who participated in the Polish General Social Survey (GSS).<sup>5</sup> The data collected on these adolescents included responses to questionnaire items about the euro currency adoption.

All computations and graphics in this paper have been done in `pOLCA` (Linzer and Lewis 2011) package of R.

The following variables (questions) in the year 2008 were used in the analysis:

- x1: Do you expect any benefits for the Polish economy (if zloty is replaced with the euro)?
- x2: Do you expect any benefits for the Polish entrepreneurs?
- x3: Do you expect any personal benefits?
- x4: Are you satisfied with the direction of the economy in our country?
- x5: Do you expect that prices are going to increase/decrease or stay the same?
- x6: Are you interested in politics?

We also analyzed the covariates:

- z1: age,
- z2: sex,
- z3: education.<sup>6</sup>

In questions 1, 2, 3, 4, 6 respondents could choose one of the following response options: definitely yes, rather yes, rather no, definitely no, I don't know.<sup>7</sup> The possible responses to the question 5 were: the prices are going to: increase (1), decrease (2), don't change (3). Because the data are rather skewed in the sense that many respondents used scores low or high the analysis is presented of dichotomized items (for negative vs. positive attitude) (see Vermunt 2010; Collins and Lanza 2011, p. 10).

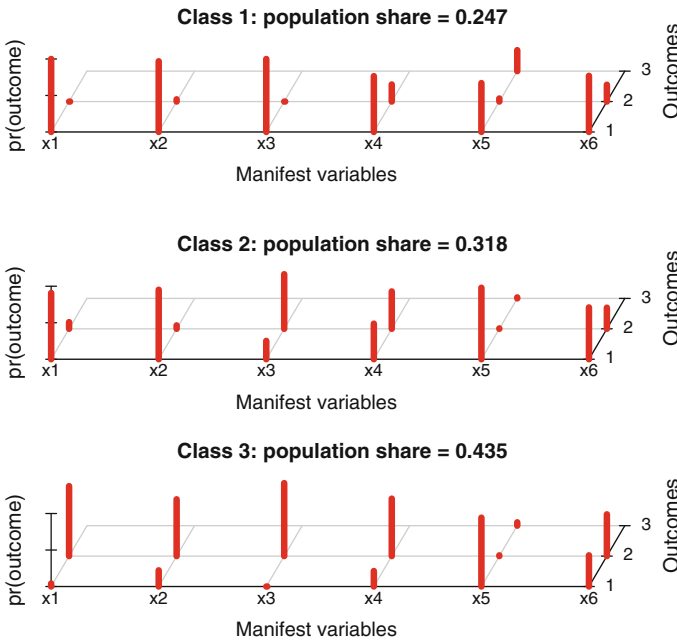
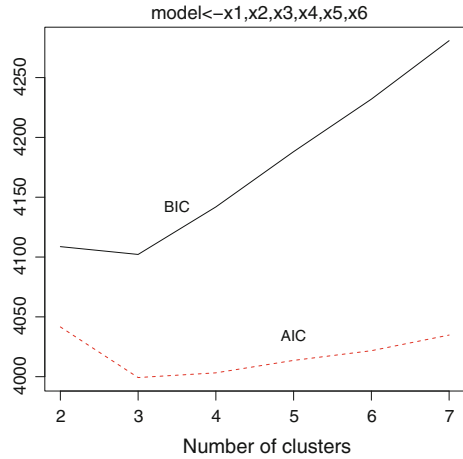
A reasonable theoretical approach might indicate that there are three latent classes of survey respondents. Euro enthusiastic, euro sceptics, and those who are more or less neutral. Supporters of euro will tend to respond favorably towards the euro, with

<sup>5</sup> The public data set, available at <http://pgss.iss.uw.edu.pl/index.php?show=wrowadzenieE/index.html&v=E>, see also (Cichomski and Jerzynski 2009).

<sup>6</sup> We considered three categories of this covariate: primary education, secondary education and tertiary education.

<sup>7</sup> The original data set contains 1,050 subjects. We analyzed the completed answers. The subjects who provided a response "I don't know" were not included.

**Fig. 2** The information criteria values

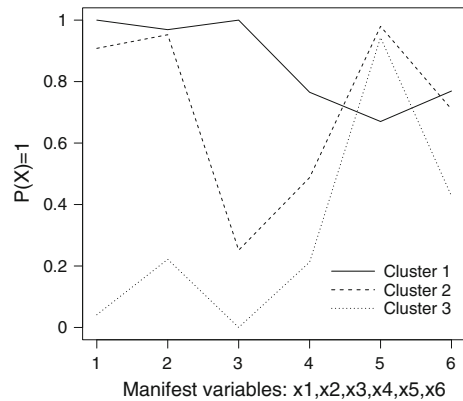


**Fig. 3** The conditional probabilities

the reverse being the case for euro sceptics. Those in neutral group will not have strong opinion about either attitude? We might further expect that falling into one of these three groups is a function of each individuals education, sex and age. We can investigate this hypothesis using a latent class model.

The optimal number of clusters was chosen using information criteria for the basic model (Collins and Lanza 2011), so we decided to choose three latent classes (see Fig. 2).

**Fig. 4** The probabilities of positive responses for three-class latent model



We estimated parameters of three components using the EM algorithm. In further analysis we ran the test for significance of the coefficients. For the three components all of the coefficients were significantly different from 0. By examining the estimated class-conditional response probabilities, we confirmed that the model identified the three groups, with 24.7 % in the pro-euro group, 43.5 % in the anti-euro group and 31.8 % in the neutral group.

We labeled the smallest latent class, which included one-fourth of the subjects, euro enthusiastic. The next largest latent class, including 31.8 % of the subjects, was labeled euro neutral. The third class, euro sceptics, contained about 43.5 % of the sample (see Fig. 3).

We examined the pattern of the probabilities of a “Yes” response to show how it was consistent with what one would expect, based on the labels chosen for the latent classes. The pattern is shown in a graphical depiction in Figs. 3 and 4.

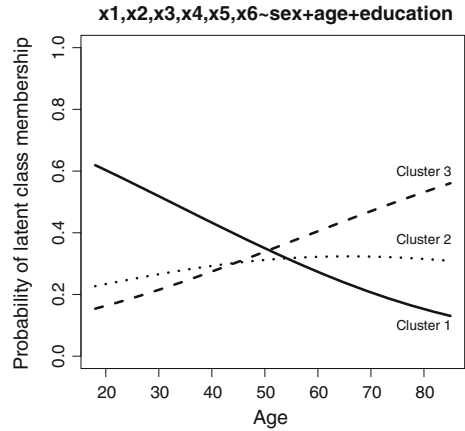
Latent class 1, euro supporters, was characterized by a high probability of positive responding to all of the variables. There was also the lowest percentage of respondents concerned about the price increases associated with the adoption of the euro.

In contrast, those in latent class 3, euro sceptics, were characterized by a low probability of positive answers to all of the variables (with an exception of the question referring to the price increases). There was 94 % of people afraid of price increases.

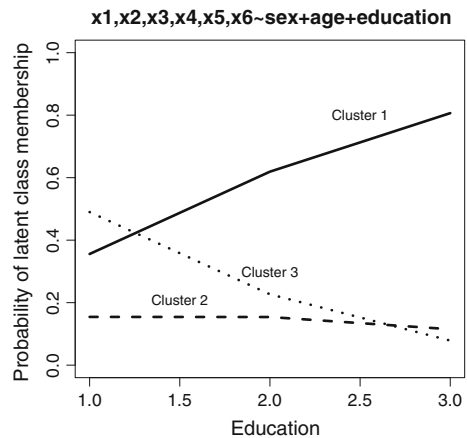
In the second class, people were not able to form an opinion. They were in favor of adopting the euro to some extent. The reasons to support euroization were that adopting the euro would bring benefits for the Polish economy and entrepreneurs, but they didn’t believe (75 % of respondents) in the improvement in their quality of the everyday life. They were interested in politics, but there was the highest percentage of citizens (98 %) who expressed the opinion that euro price conversion would lead to increased prices.<sup>8</sup>

<sup>8</sup> We estimated also parameters of the latent class model without covariates. The model identified the three groups, with 19.3 % in the pro-euro group, 44.0 % in the anti-euro group and 36.7 % in the neutral group. The pattern of the probabilities of a “Yes” responses (the estimated class-conditional response probabilities) were quite similar to those for the latent class model with covariates (with the biggest difference of 12 % for x3 variable in the euro neutral group—63 % of respondents didn’t believe in the improvement in their quality of the everyday life).

**Fig. 5** Predicted prior probabilities of latent class membership at ages 18–85



**Fig. 6** Predicted prior probabilities of latent class membership at varying levels of education



We were interested in whether the effect of sex modified the effect of education and age on the respondents' approach to the euro. We specified the interaction model and examined the interaction effects. The hypothesis test showed that there was no interaction statistically significant, but separate variables age, gender and education had an influence on the prior probabilities of class membership.

To interpret the estimated generalized logit coefficients of covariates, we calculated and plotted the prior probabilities of class membership, at varying levels of age. In Figs. 5 and 6 the estimated prior probabilities of class membership are given separately for each education level and every age.<sup>9</sup>

We would support the view that older people would be more critical towards the euro as they might find the adjustment to the new currency more difficult than younger people. Furthermore, elderly people in Poland remember currency changeovers as

<sup>9</sup> The figures at varying levels of one covariates were prepared for the most frequent category for the all covariates (see Grün and Leisch 2008, p. 21; Linzer and Lewis 2011, p. 26; Wittek 2011 pp. 236–239).

poverty-inducing and they may be particularly sensitive to all potential aspects of sovereignty issues due to the specific Polish historical conditions.

Young people (at age of 20) have over a 60 % prior of belonging to the euro enthusiastic group, while old people have almost the same (60 %) prior of belonging to the euro-sceptic class. The prior probability of belonging to the neutral group, is at the similar level in every age.

As expected, well educated people are very unlikely to belong to the third class. Respondents with higher education level have over a 80 % prior of belonging to the supporters of euro group. However it is interesting to observe that regardless of education level, respondents are very likely to belong to the neutral group.

Due to the space limit we, do not present the figure for the gender covariate. The probability of belonging to the first class is significantly higher for men than for women (62 % for men and 29 % for women), while women have slightly higher prior of belonging to the neutral group and 23 % higher (than men) to the third group.

## 6 Conclusions

We analyzed the Polish opinions on the euro adoption. We applied latent class analysis for grouping and detecting inhomogeneities of Polish attitude toward the euro currency for the data collected as a part of the Polish General Social Survey using software of R. By examining the estimated class-conditional response probabilities, we confirmed that the society can be divided into three groups—euro supporters, euro sceptics and euro neutral. We also showed the influence of covariates on the prior probabilities (Figs. 5, 6).

In future research it would be worthwhile to compare the results with the three-step analysis of covariates estimation (Vermunt 2010). The analysis should also be extended to the data at different points of time (i.e. 2008, 2010, 2012). The longitudinal data would make it possible to examine transition in Poles' attitude over time (latent transition analysis).

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Agresti A (2002) *Categorical data analysis*. John Wiley and Sons, Hoboken
- Bandeen-Roche K, Miglioretti DL, Zeger SL, Rathouz PJ (1997) Latent variable regression for multiple discrete outcomes. *J Am Stat Assoc* 92(40):123–135
- Banfield JD, Raftery AE (1993) Model-based Gaussian and non-Gaussian clustering. *Biometrics* 49(3): 803–821
- Biernacki C, Celeux G, Govaert G (1999) Choosing models in model-based clustering and discriminant analysis. *J Stat Comput Simul* 64:49–71
- Cichowski B, Jerzynski T (2009) *Polskie Generalne Sondaze Spoleczne: skumulowany komputerowy zbiór danych 1992–2008*. Instytut Studiów Społecznych, Uniwersytet Warszawski, Warszawa
- Collins LM, Lanza ST (2011) *Latent class and latent transition analysis with applications in the social, behavioral, and health sciences*. John Wiley and Sons, Wiley, pp 151–177

- Clogg CC (1981) New developments in latent structure analysis. In: Jackson DJ, Borgotta EF (eds) *Factor analysis and measurement in sociological research*. Sage Publications, Beverly Hills, pp 215–246
- Dayton CM, Macready GB (1988) Concomitant-variable latent-class models. *J Am Stat Assoc* 83(401): 173–178
- Dempster AP, Laird NP, Rubin DB (1977) Maximum likelihood for incomplete data via the EM algorithm (with discussion). *J R Stat Soc* 39:1–38
- Fraley C, Raftery AE (2002) Model-based clustering, discriminant analysis, and density estimation. *J Am Stat Assoc* 97:611–631
- Grün B, Leisch F (2008) FlexMix Version 2: finite mixtures with concomitant variables and varying and constant parameters. *J Stat Softw* 28(4):1–35
- Hagenaars AJ, McCutcheon AL (2002) *Applied latent class analysis*. Cambridge University Press, Cambridge
- Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90:773–795
- Keribin C (1998) Consistent estimate of the order of mixture models. *Comptes Rendus de l'Academie des Sciences, Serie I-Mathematicues* 326:243–248
- Lazarsfeld PF (1950a) The logical and mathematical foundations of latent structure analysis. In: Stouffer SA (ed) *Measurement and prediction, the American soldier: studies in social psychology in World War II*. Princeton University Press, Princeton, pp 362–412
- Lazarsfeld PF (1950b) The interpretation and computation of some latent structures. In: Stouffer (ed) *Measurement and prediction, the American soldier: studies in social psychology in World War II*. Princeton University Press, Princeton, pp 413–472
- Lazarsfeld PF, Henry NW (1968) *Latent structure analysis*. Houghton Mifflin, Boston
- Linzer DA, Lewis J (2011) poLCA: an R package for polytomous variable latent class analysis. *J Stat Softw* 42(10):1–29
- McLachlan GJ, Peel D (2000) *Finite mixture models*. Wiley, New York
- Meilijson I (1989) A fast improvement to the EM algorithm on its own terms. *J R Stat Soc* 51(1):127–138
- Osinska J, Toroj A (2012) Greek ricochet? What drove Poles' attitudes to the euro in 2009–2010. *Bank i Kredyt* 43(4):29–84
- Stanford D, Raftery AE (2000) Principal curve clustering with noise. *IEEE Trans Pattern Anal Mach Intell* 22:601–609
- Titterton DM, Smith AFM, Makov UE (1985) *Statistical analysis of finite mixtures of distributions*. Wiley, New York
- Vermunt JK (1997) *Log-linear models for event histories, advanced quantitative techniques in the social sciences series*. Sage Publikations, Thousand Oaks
- Vermunt JK, Magidson J (2002) Latent class cluster analysis. In: Hagenaars J, McCutcheon A (eds) *Applied latent class analysis*. Cambridge University Press, Cambridge, pp 89–106
- Vermunt JK (2010) Latent class modeling with covariates: two improved three-step approaches. *Political Anal* 18:450–469
- Witek E (2010) Analysis of massive emigration from Poland—the model-based clustering approach. In: *Proceedings of the 32nd annual conference of the Gesellschaft für Klassifikation*. Springer, Berlin, pp 615–624
- Witek E (2011a) Modele mieszanek dla danych jakosciowych. In: Gatnar E, Walesiak M (eds) *Analiza danych jakosciowych i symbolicznych z wykorzystaniem programu R*. C. H. Beck, Warszawa
- Witek E (2011b) The comparison of model-based clustering with heuristic clustering methods. In: Domanski Cz, Bialek J (eds) *Folia Oeconomica* 255, *Methodological aspects of multivariate statistical analysis, statistical models and applications*. Wydawnictwo Uniwersytetu Lodzkiego, Lodz, pp 191–197
- Wolfe JH (1963) *Object cluster analysis of social areas*. University of California, Barkeley Master's thesis