Fall 1996, Vol. 21, No. 3, pp. 215–229

Estimating the Concomitant-Variable Latent-Class Model With the EM Algorithm

Peter G. M. van der Heijden and **Jos Dessens**
University of Utrecht
Ulf Bockenholt
University of Illinois at Urbana-Champaign

Key words: *categorical data, data analysis, EM algorithm, latent class analysis, latent variables*

Latent class analysis assumes the existence of a categorical latent variable that explains the relations between a set of categorical manifest variables. Simultaneous latent class analysis deals with sets of multiway contingency tables simultaneously. In this way an explanatory categorical grouping variable is related to latent class results. In this article we discuss a tool called the concomitant-variable latent-class model, which generalizes this work to continuous explanatory variables. An EM estimation procedure to estimate the model is worked out in detail, and the model is applied to an example on juvenile delinquency.

Latent class analysis (LCA; see, for example, Goodman, 1974; Langeheine, 1988) assumes the existence of a categorical latent variable that explains the relations between a set of categorical manifest variables. Given the level of this unknown (i.e., latent) variable, the manifest variables are independent. Given the latent class, conditional probabilities specifying the levels of the manifest variables describe the relation between the manifest variables and the latent variable. Latent class probabilities specify the probability that an observation falls into each of the levels of the latent variable.

Simultaneous LCA (Clogg & Goodman, 1984) is the application of LCA to a set of multiway contingency tables simultaneously. Such a set may be defined by a grouping variable, such as different years or different sexes. Homogeneity constraints across the groups can be imposed in order to test whether a simplified interpretation is allowed. Further extensions are provided by Dayton and Macready (1988, 1989), Hagenaars (1990), van der Heijden, Mooijaart, and de Leeuw (1992), and Formann (1992). Hagenaars shows that if one relates more than one explanatory categorical variable to the latent class probabilities, it is possible to restrict these probabilities by loglinear models. This line is also taken up by van der Heijden et al., who restrict

---

We gratefully acknowledge the helpful comments of C. M. Dayton, A. K. Formann, J. A. Hagenaars, F. van der Pol, and an anonymous reviewer on an earlier version of this article. This research was partially supported by NSF Grant SBR-9409531.

what they call latent budget analysis in such a way that a version of simultaneous LCA is obtained, where categorical or grouped continuous variables are related to the latent class probabilities by a multinomial logit model (cf. McCutcheon, 1994). Formann presents a generalization where not only the latent class probabilities but also the conditional probabilities are restricted by multinomial logit models.

Dayton and Macready (1988, 1989) provided an extension of simultaneous LCA by allowing explanatory variables to be continuous, that is, quantitative and possibly different for each observation. To the best of our knowledge, this model has seldom been used, although our impression is that it deserves further study. In a recent review paper, Clogg (1995) describes this as an interesting research area. Similarly, DeSarbo and Wedel (1994) review recent developments in this and closely related areas and also give a survey of applications. In the present article we present an estimation procedure for this model that differs from the procedure employed by Dayton and Macready (1988, 1989), and subsequently discuss an example.

Latent Class Analysis

Let there be three categorical manifest variables having categories indexed by u ($u = 1, \ldots, U$), v ($v = 1, \ldots, V$), and w ($w = 1, \ldots, W$), respectively. LCA assumes a categorical latent variable with categories indexed by x ($x = 1, \ldots, X$), such that given a level x of this latent variable, the manifest variables are independent. This can be formalized as follows. We denote the joint (latent) probability of level x of the latent variable and response pattern u, v, and w by π_{uvwx}. Let π_x be the probability of falling into latent class x, and let $\pi_{u|x}$, $\pi_{v|x}$, and $\pi_{w|x}$ be conditional probabilities of falling into levels u, v, and w, respectively, given x. Then

$$\pi_{uvwx} = \pi_x \pi_{u|x} \pi_{v|x} \pi_{w|x}, \qquad \forall u, v, w, x, \tag{1}$$

with restrictions

$$\sum_{x=1}^{X} \pi_x = 1 \text{ and } \sum_{u=1}^{U} \pi_{u|x} = \sum_{v=1}^{V} \pi_{v|x} = \sum_{w=1}^{W} \pi_{w|x} = 1, \forall x. \tag{2}$$

The unobserved probabilities are related to the expected probabilities π_{uvw} by

$$\pi_{uvw} = \sum_{x=1}^{X} \pi_{uvwx}. \tag{3}$$

In practical applications, LCA is often used as a tool for data reduction. After a latent class analysis, conditional probabilities $\pi_{x|uvw} = \pi_{uvwx}/\Sigma_x \pi_{uvwx}$ can be derived and used to assign every observation having response pattern u, v, w to that latent class x ($x = 1, \ldots, X$) for which $\pi_{x|uvw}$ is highest. These

class assignments are then related to external variables. Usually the class assignment is done by the modal rule; that is, an observation having response pattern u, v, w is classified in latent class z if for $z = x$ the conditional probability $\pi_{z|uvw}$ is highest. Such a data analysis procedure is suboptimal. Firstly, the classification will often lead to considerable errors. If an observation is classified into class z, the error is $1 - \pi_{x|uvw}$. For example, as Hagenaars (1993) points out,

> most probably, several individuals will be assigned to the wrong latent class, with the result, among other things, that the true relations between latent and external variables are not correctly reflected by the relations between the "observed" latent class scores and external variables. (p. 33)

Secondly, it is preferable to have one model or data analysis procedure that performs both steps not subsequently but simultaneously. In this way, hypotheses concerning relations between explanatory variables and latent-class parameters can be tested directly.

Such a model was proposed by Dayton and Macready (1988, 1989) as a submodel of what they called the *concomitant-variable latent-class model.* This model resolves both problems and has the added advantage that in situations where the ordinary latent class model is not identified or has zero degrees of freedom, this submodel of the concomitant-variable latent-class model can have a positive number of degrees of freedom, dependent on the number of concomitant variables.

The Concomitant-Variable Latent-Class Model

In the concomitant-variable latent-class model, the latent class probabilities are related to explanatory variables. The model is defined as follows. Let there be n observations, indexed by i ($i = 1, \ldots, n$). For each observation, there are M explanatory variables, indexed by m. They may be grouped or ungrouped continuous variables, or dummy variables. These explanatory variables are collected into a matrix $\mathbf{Z}$ of order $n \times M$, with elements z_{im}. The concomitant-variable latent-class model is defined by

$$\pi_{uvwx|i} = \pi_{x|i} \pi_{u|x} \pi_{v|x} \pi_{w|x}, \tag{4}$$

where $\pi_{uvwx|i}$ is the conditional probability that observation i falls into u, v, w and level x of the latent variable, and $\pi_{x|i}$ is the conditional probability that observation i falls into level x of the latent variable; $\pi_{x|i}$ depends on the scores of observation i on the M explanatory variables. Notice that $\Sigma_u \Sigma_v \Sigma_w \Sigma_x \pi_{uvwx|i} = 1$ and $\Sigma_x \pi_{x|i} = 1$. Similar to (3),

$$\pi_{uvw|i} = \sum_{x=1}^{X} \pi_{uvwx|i} \tag{5}$$

so that $\Sigma_u \Sigma_v \Sigma_w \pi_{uvw|i} = 1$. A submodel of the concomitant-variable latent-class model is obtained by relating the explanatory variables in **Z** to $\pi_{x|i}$ by a multinomial logit model:

$$\pi_{x|i} = \frac{\exp\left(\sum_{m=1}^{M} z_{im}\gamma_{mx}\right)}{\sum_{x=1}^{X} \exp\left(\sum_{m=1}^{M} z_{im}\gamma_{mx}\right)}, \quad (6)$$

where γ_{mx} is a parameter relating explanatory variable m to latent class x. In order to identify the parameters, the restriction $\gamma_{m1} = 0$ is used.

If the variables in **Z** are dummy variables—for example, two dummy variables for religion with the three categories Catholic, protestant, and no religion—then concomitant-variable LCA is equivalent to simultaneous LCA where the conditional probabilities $\pi_{u|x}$, $\pi_{v|x}$, and $\pi_{w|x}$ must be homogeneous over the three groups, but where the latent class probabilities may differ over the three groups.

A further extension of this model is obtained when multinomial logit constraints are imposed on the conditional probabilities (cf. van der Heijden et al., 1992, and Formann, 1992, who already proposed this for explanatory categorical variables). Thus the model becomes

$$\pi_{uvwx|i} = \pi_{x|i}\pi_{u|xi}\pi_{v|xi}\pi_{w|xi}, \quad (7)$$

where $\Sigma_u \pi_{u|xi} = 1$ for all x and i, and with similar constraints for $\pi_{v|xi}$ and $\pi_{w|xi}$. For each x, a multinomial logit model can be defined for $\pi_{u|xi}$ (and, similarly, for $\pi_{v|xi}$ and $\pi_{w|xi}$), so

$$\pi_{u|xi} = \frac{\exp\left(\sum_{m=1}^{M} z_{im}\gamma_{mux}\right)}{\sum_{u=1}^{U} \exp\left(\sum_{s=1}^{M} z_{is}\gamma_{sux}\right)}, \quad (8)$$

where γ_{mux} is a parameter relating explanatory variable m to category u for latent class x. In order to identify the parameters, the restriction $\gamma_{m1x} = 0$ is used.

If the variables in **Z** are dummy variables—for example, two dummy variables for religion with the three categories Catholic, protestant, and no religion—then restrictions like the one in (8) allow for heterogeneity of conditional probabilities in simultaneous LCA.

Estimation

We derive maximum likelihood estimates of the parameters by using the EM algorithm (Dempster, Laird, & Rubin, 1977). This differs from Dayton

and Macready (1988), who used the simplex method: a grid search technique over the entire parameter space. We first show how the EM algorithm works for ordinary LCA (cf. Mooijaart & van der Heijden, 1992), after which its extension to concomitant-variable LCA is easy.

The estimation problem of the latent class model can be conceived of as an estimation problem with missing data: the observations on the latent variable are missing. Then, in the EM algorithm, the likelihood for the complete data (i.e., both the data that are observed and the data that are unobserved) is specified. Assuming a multinomial distribution, the loglikelihood is

$$\log L = \sum_{u,v,w,x}^{U,V,W,X} n_{uvwx} \log \pi_{uvwx} + \text{constant}, \quad (9)$$

where both n_{uvwx} and the parameter estimates are unknown. The EM algorithm consists of two steps. In the E-step, expectations of the missing data n_{uvwx} are found given the observed data n_{uvw} and the parameter estimates:

$$\bar{n}_{uvwx} = n_{uvw} \frac{\pi_{uvwx}}{\pi_{uvw}}, \quad (10)$$

where π_{uvwx} is calculated via (1) using the current parameter estimates. In (10) the observed frequency n_{uvw} is distributed over the X latent classes.

In the M-step (M for maximization), the loglikelihood is maximized over the parameters, using the updates derived in (10). The loglikelihood function becomes

$$f(\pi_x, \pi_{u|x}, \pi_{v|x}, \pi_{w|x}, \gamma, \delta_{ux}, \delta_{vx}, \delta_{wx}) = \sum_{u,v,w,x}^{U,V,W,X} \bar{n}_{uvwx} \log(\pi_x \pi_{u|x} \pi_{v|x} \pi_{w|x}) -$$

$$\gamma\left[\left(\sum_{x=1}^{X} \pi_x\right) - 1\right] - \sum_{x=1}^{X}\left\{\delta_{ux}\left[\left(\sum_{u=1}^{U} \pi_{u|x}\right) - 1\right]\right\} - \sum_{x=1}^{X}\left\{\delta_{vx}\left[\left(\sum_{v=1}^{V} \pi_{v|x}\right) - 1\right]\right\} -$$

$$\sum_{x=1}^{X}\left\{\delta_{wx}\left[\left(\sum_{w=1}^{W} \pi_{w|x}\right) - 1\right]\right\} + \text{constant}, \quad (11)$$

where γ, δ_{ux}, δ_{vx}, and δ_{wx} are Lagrange multipliers accounting for the restrictions in (2). This likelihood (11) can be split into parts that can be optimized independently, that is,

$$f(\pi_x, \pi_{u|x}, \pi_{v|x}, \pi_{w|x}, \gamma, \delta_{ux}, \delta_{vx}, \delta_{wx}) = \sum_{x}^{X} \bar{n}_{+++x} \log \pi_x + \sum_{u,x}^{U,X} \bar{n}_{u++x} \log \pi_{u|x} +$$

$$\sum_{v,x}^{V,X} \bar{n}_{+v+x} \log \pi_{v|x} + \sum_{w,x}^{W,X} \bar{n}_{++wx} \log \pi_{w|x} -$$

$$\gamma\left[\left(\sum_{x=1}^{X} \pi_x\right) - 1\right] - \sum_{x=1}^{X}\left\{\delta_{ux}\left[\left(\sum_{u=1}^{U} \pi_{u|x}\right) - 1\right]\right\}$$
$$- \sum_{x=1}^{X}\left\{\delta_{vx}\left[\left(\sum_{v=1}^{V} \pi_{v|x}\right) - 1\right]\right\} - \sum_{x=1}^{X}\left\{\delta_{wx}\left[\left(\sum_{w=1}^{W} \pi_{w|x}\right) - 1\right]\right\} + \text{constant.} \quad (12)$$

Thus, if we want to maximize (12) over π_x, we should maximize

$$f^*(\pi_x, \gamma) = \sum_{x}^{X} \bar{n}_{+++x} \log \pi_x - \gamma\left[\left(\sum_{x=1}^{X} \pi_x\right) - 1\right]. \quad (13)$$

Then by taking derivatives, making them equal to zero, and solving for γ, we get

$$\frac{\partial f(\pi_x, \gamma)}{\partial \pi_x} = \frac{\bar{n}_{+++x}}{\pi_x} - \gamma = 0 \Leftrightarrow \pi_x = \frac{\bar{n}_{+++x}}{\gamma}, \quad (14)$$

and by using $\Sigma_x \pi_x = 1$ to solve for γ, we find $\bar{\pi}_x = \bar{n}_{+++x}/\bar{n}_{++++}$. In a similar way, we find

$$\bar{\pi}_{u|x} = \frac{\bar{n}_{u++x}}{\bar{n}_{u+++}}; \; \bar{\pi}_{v|x} = \frac{\bar{n}_{+v+x}}{\bar{n}_{+v++}}; \text{ and } \bar{\pi}_{w|x} = \frac{\bar{n}_{++wx}}{\bar{n}_{++w+}}. \quad (15)$$

We will now derive the estimates for the concomitant-variable model as described in (4), (5), and (6). The loglikelihood is

$$\log L = \sum_{u,v,w,x,i}^{U,V,W,X,n} n_{uvwi} \log \pi_{uvwxi} + \text{constant.} \quad (16)$$

The E-step is simpler than in (10), because each observation is allowed to be unique with respect to its explanatory variables, and therefore $n_{uvwi} = 1$:

$$\bar{n}_{uvwx|i} = n_{uvwi} \frac{\pi_{uvwx|i}}{\pi_{uvw|i}} = \frac{\pi_{uvwx|i}}{\pi_{uvw|i}}. \quad (17)$$

In the M-step, the loglikelihood can be split into parts that can be optimized independently. The part for $\pi_{x|i}$ is

$$f(\gamma_{mx}) = \sum_{x,i}^{X,n} \bar{n}_{+++xi} \log\left(\frac{\exp\left(\sum_{m=1}^{M} z_{im}\gamma_{mx}\right)}{\sum_{x=1}^{X} \exp\left(\sum_{m=1}^{M} z_{im}\gamma_{mx}\right)}\right). \quad (18)$$

so that in each M-step a multinomial logit model must be fitted to the values n_{+++xi}. We fit this multinomial logit model using the Newton-Raphson method (see Bock, 1975, for details; Maddala, 1983). The other parameters can be derived in ways analogous to that described in (17) and (18); for example, $\bar{\pi}_{u|x} = \bar{n}_{u++x+}/\bar{n}_{u++++}$.

For the extension of the concomitant-variable latent-class model described in (7) and (8), the EM algorithm is easily adapted. For the E-step, (17) can still be used. For the M-step, the likelihood can still be split up in parts that can be estimated independently. Imposing additional restrictions as in (8) does not influence the way in which $\pi_{x|i}$ has to be estimated. But where $\pi_{u|x}$ could be estimated easily by $\bar{\pi}_{u|x} = \bar{\pi}_{u++x+}/\bar{\pi}_{u++++}$, now $\pi_{u|xi}$ has to be estimated by optimizing the following part of the likelihood over γ_{mux}:

$$f(\gamma_{mux}) = \sum_{u,x,i}^{U,X,n} \bar{n}_{u++xi} \log\left(\frac{\exp\left(\sum_{m=1}^{M} z_{im}\gamma_{mux}\right)}{\sum_{u=1}^{U} \exp\left(\sum_{s=1}^{M} z_{is}\gamma_{sux}\right)}\right). \quad (19)$$

This can be maximized by fitting a multinomial logit model to n_{u++xi} for each x separately.

The EM algorithm converges gradually but surely; however, it may converge to a local maximum of the likelihood. Different initial estimates should be essayed to increase confidence that the maximum found is indeed the global maximum (see Formann, 1992, for more details). The EM algorithm is numerically very stable for the concomitant-variable latent-class model described in (4), (5), and (6). However, for the extension described in (7) and (8), the algorithm may prove to be rather unstable. In particular, it may not be possible to identify subsets of multinomial logit parameters (19) when the number of observations in combinations of u and x is small. Instead, it may be necessary to impose additional constraints that apply to several classes simultaneously. Currently, we are working on a general approach of imposing such constraints to simultaneously model the effects of concomitant variables on $\pi_{x|i}$ and $\pi_{u|xi}$.

Because the explanatory variables are continuous, the table is sparse. Therefore, a model cannot be tested against the data as in ordinary LCA. However, given a choice for a specific number of latent classes X, it is possible to compare the loglikelihoods of two nested models by using 2 times the difference in the loglikelihoods of two nested models and using the difference in the number of parameters as the number of degrees of freedom. If the explanatory variables are dummy variables, then the model reduces to the simultaneous LCA of Clogg and Goodman (1984), or restricted versions of Hagenaars (1990), and tests of the model can be carried out against the data if the observed frequencies are sufficiently large.

Example: Crime Among Four Ethnic Groups

The Netherlands Ministry of Justice investigated the differences in involvement in crime among youth from four ethnic groups: Moroccans, Turks, Surinamese, and Dutch. To control for the generally lower socioeconomic status of the first three ethnic groups, the Dutch sample consisted of youngsters who lived on the same streets as the youngsters from the other ethnic groups. The total sample size was 788. For more details, see Junger (1990). Among other things, three crime measures were gathered from the police registration: property crime, aggression against persons, and vandalism. In van der Heijden et al. (1992), the age of the youngsters was coded into three categories: 12–13, 14–15, and 16–17. The authors' research questions were: (a) Are there two or more types of crime? (b) Are these types of crime the same for each of the ethnic groups? (c) If so, how is group membership (i.e., ethnic group and age group membership) related to these types of crime? The answers found by latent budget analysis were: (a) There is no evidence for more than two types of crime. (b) There is no evidence for different types for different groups. (c) There is a relation between ethnic group and type of crime, as well as a linear age effect on type of crime.

In this article, we resume the analysis of van der Heijden et al. (1992). This is justified because their latent budget analysis is a special case of the concomitant-variable latent-class model where the explanatory variables are grouped continuous variables (age) and dummy variables (ethnic group). Here six additional explanatory continuous variables are studied, all having to do with family integration. Social control theory (Hirschi, 1969) predicts that a child who has a better integrated family will be less inclined to commit crime. The six variables are scales derived by means of principal component analysis and a reliability analysis. They are *direct control*, *family arguments*, *violence*, *emotional bonds*, *importance of school*, and *general evaluation of the family*.[1]

We begin with a search for an adequate model (see Table 1). The first model (M_0) assumes that the parameter estimates for $\pi_{x|i}$ are the same for each i. This model has a loglikelihood of −596.6. The second model, M_1, introduces two dummy variables for age. The loglikelihood of this model is −584.9, and the difference with M_0 is significant: $G^2(M_1 - M_0) = 23.2$, $df = 2$. In M_3, three dummy variables for ethnicity are added. The likelihood then increases to −575.9. As in the analysis of van der Heijden et al. (1992), we assume in M_3 that the effect of age is linear, and this hypothesis is not rejected ($G^2(M_3 - M_2) = 0.0$, $df = 1$). We will now add continuous variables. In M_4, the six variables are added, and the likelihood is increased considerably to −559.4 ($G^2(M_4 - M_3) = 33.1$, $df = 6$). It is now interesting to ask whether these family integration variables can be seen as explaining the role of ethnicity, but this is not the case. We find that the difference between M_4 and M_5 is significant. From models M_6 through M_9 it follows that the variables *importance of school* (M_6), *emotional bond* (M_7), and *violence* (M_8) can be dropped from the model. However, it is not permissible to drop any of the remaining variables; for example, the difference between M_8 and M_9 is significant, and *family arguments* should therefore be included in the model.

We will now discuss parameter estimates of models M_2 and M_8. The parameter estimates for M_2 (see Tables 2–4) are very similar to those found in van der Heijden et al.'s (1992) latent budget analysis. The conditional proportions (Table 2) indicate that Latent Class 1 is the delinquent latent class. Here the probabilities of being registered for property crime, aggression against persons, and vandalism are .931, .199, and .222, respectively. In the

TABLE 1
Fit of models

Model	Explanatory variables	Loglikelihood	Compared to	*df*	G^2 diff.
M_0	Mean	−596.6			
M_1	Age	−584.9	M_0	2	23.2
M_2	Age, ethnicity	−575.9	M_1	3	18.1
M_3	Age linear, ethnicity	−575.9	M_2	1	0.0
M_4	Age linear, ethnicity, DC, EB, FA, V, ISB, GE	−559.4	M_3	6	33.1
M_5	Age linear, DC, EB FA, V, ISB, GE	−570.8	M_4	3	23.0
M_6	Age linear, ethnicity, DC, EB, FA, V, GE	−559.4	M_4	1	0.0
M_7	Age linear, ethnicity, DC, FA, V, GE	−559.5	M_6	1	0.2
M_8	Age linear, ethnicity, DC, FA, GE	−560.9	M_7	1	2.8
M_9	Age linear, ethnicity, DC, GE	−563.6	M_8	1	5.5

Note. DC = direct control, EB = emotional bond, FA = family arguments, V = violence, ISB = importance of school, GE = general evaluation, G^2 diff. = −2 times the difference between the loglikelihoods of the two models.

TABLE 2
Parameter estimates of M_2*: Estimates for the conditional probabilities* $\pi_{u|x}$ *(property crime),* $\pi_{v|x}$ *(aggression), and* $\pi_{w|x}$ *(vandalism)*

	Property crime		Aggression		Vandalism	
Latent class (x)	No	Yes	No	Yes	No	Yes
1	.069	.931	.801	.199	.778	.222
2	.949	.051	.992	.008	.975	.025

TABLE 3

Parameter estimates of M_2*: Estimates for the multinomial logit parameters* γ_{mx} *(with standard errors in parentheses)*

Variable (*m*)	Latent class (*x*) 1	2
Mean	0	2.24 (0.80)
Age 12–13	0	1.71 (0.57)
Age 14–15	0	0.80 (0.13)
Moroccan	0	−1.83 (0.50)
Turkish	0	−0.87 (0.53)
Surinamese	0	−1.02 (0.57)

Note. In order to identify the parameters, the restriction $\gamma_{m1} = 0$ is used.

nondelinquent Latent Class 2 these probabilities are .051, .008, and .025. The multinomial logit parameters (Table 3) for the first latent class are restricted to zero. Estimates of probabilities $\pi_{x|i}$ (Table 4) can be derived as in the following example: For Moroccan boys of ages 12 and 13, the probability of falling in Latent Class 1 is [exp 0]/[exp 0 + exp(2.24 + 1.71 − 1.83)] = .107, and the probability of falling in Latent Class 2 is [exp(2.24 + 1.71 − 1.83)]/[exp 0 + exp(2.24 + 1.71 − 1.83)] = .893. For ease of interpretation we have specified these estimates for all 12 groups. This shows that for each ethnic group the proportion of delinquents increases considerably with age. Moroccans have the highest level of delinquence, followed by Turks and Surinamese, followed by Dutch.

TABLE 4

Parameter estimates of M_2*: Estimates of* $\pi_{x|i}$ *made using Equation 6 for each of the 12 age-ethnicity groups*

Ethnicity	Age	Latent class (*x*) 1	2
Moroccan	12–13	.107	.893
	14–15	.230	.770
	16–17	.398	.602
Turkish	12–13	.044	.956
	14–15	.102	.897
	16–17	.203	.797
Surinamese	12–13	.051	.949
	14–15	.117	.883
	16–17	.228	.772
Dutch	12–13	.019	.981
	14–15	.046	.954
	16–17	.096	.904

In Tables 5–7, we find the parameter estimates for M_8. Again, the first latent class is the delinquent latent class, with conditional probabilities (Table 5) similar to those for M_2, except for $\pi_{u|x}$, which is now lower (.857). In Table 7, the parameter estimates $\pi_{x|i}$ are given only for $x = 1$. In Column A we find the estimates for $\pi_{x|i}$ for the 12 groups when the variables *direct control*, *family arguments*, and *general evaluation* are at their means. In Columns B, C, and D we see what happens when each of them is 1 standard deviation above its mean. In Column E, results are displayed for the situation in which the three variables are simultaneously 1 standard error above their means. Children with a high degree of direct control by their parents, few family arguments, and a positive general evaluation have a rather low probability of becoming delinquent, with the exception of Moroccans, who still have a probability of .15 for boys aged 16–17.

A graphical display of Table 7 is provided in Figure 1. It shows three important effects: (a) The probability of being involved with crime increases with age for each ethnic group. (b) The probability of being involved with crime is highest for Moroccans, much lower for Turks and Surinamese, and lowest for Dutch boys. (c) The probability of being involved with crime

TABLE 5

Parameter estimates of M_8*: Estimates for the conditional probabilities* $\pi_{u|x}$ *(property crime),* $\pi_{v|x}$ *(aggression), and* $\pi_{w|x}$ *(vandalism)*

	Property crime		Aggression		Vandalism	
Latent class (*x*)	No	Yes	No	Yes	No	Yes
1	.143	.857	.781	.219	.752	.248
2	.935	.065	.994	.006	.978	.022

TABLE 6

Parameter estimates of M_8*: Estimates for the multinomial logit parameters* γ_{mx} *(with standard errors in parentheses)*

Variable (*m*)	Latent class (*x*) 1	2
Mean	0	9.18 (2.40)
Age linear	0	−0.75 (0.19)
Moroccan	0	−2.17 (0.46)
Turkish	0	−0.82 (0.37)
Surinamese	0	−1.07 (0.35)
Direct control	0	−0.72 (0.14)
Family arguments	0	−0.52 (0.16)
General evaluation	0	−0.63 (0.36)

Note. In order to identify the parameters, the restriction $\gamma_{m1} = 0$ is used.

TABLE 7

Parameter estimates of M_8*: Estimates of* $\pi_{1|i}$ *(the probability of falling in the delinquent Latent Class 1) made using Equation 6 for each of the 12 age-ethnicity groups when the variables* direct control, family arguments, *and* general evaluation *are manipulated*

Ethnicity	Age	A	B	C	D	E
Moroccan	12–13	.119	.075	.083	.090	.038
	14–15	.223	.147	.161	.173	.078
	16–17	.377	.267	.289	.307	.152
Turkish	12–13	.034	.021	.023	.025	.010
	14–15	.069	.043	.047	.051	.021
	16–17	.135	.086	.095	.103	.044
Surinamese	12–13	.043	.026	.029	.032	.013
	14–15	.087	.054	.060	.065	.027
	16–17	.168	.108	.108	.128	.056
Dutch	12–13	.015	.009	.010	.011	.005
	14–15	.032	.019	.022	.023	.010
	16–17	.065	.040	.045	.048	.020

Note. In Column A, the explanatory variables *direct control, family arguments*, and *general evaluation* are at their means. In Column B, *direct control* is 1 standard deviation above its mean. In Column C, *family arguments* is 1 standard deviation above its mean. In Column D, *general evaluation* is 1 standard deviation above its mean. In Column E, all three variables are 1 standard deviation above their means.

decreases if the family situation as measured by *direct control, family arguments*, and *general evaluation* is better.

Conclusion

The concomitant-variable latent-class model seems to be a useful tool for data analysis. Its results are not too difficult to interpret if we translate the multinomial logit parameters into parameters $\pi_{x|i}$. Theoretical advantages of the model are threefold. First, the problem of making errors in the process of classifying observations to latent classes is circumvented. Second, it is more elegant to work with a procedure that has one measure of model fit than with two different procedures that have their own measures of model fit. And, third, it is possible to investigate models that have zero degrees of freedom or are unidentified in ordinary LCA by using covariates in concomitant-variable LCA.

Note

[1]The indexes are derived from the following variables: (a) *direct control* (1 = high to 4.7 = low), assessed by "If you go out, do your parents know with whom?" "Do your parents tell you what time to come home (when you go out)?" "Do you listen to them?" "Do your parents ask if you have finished your homework?" "Do you have to finish you homework before going outside?" and "I can do whatever I like"; (b) *family arguments* (1 = few to 5 = many), assessed by "arguments in the family," "arguments with my father," and "arguments with my mother"; (c) *violence* (1 = no to 5 = much), assessed by how often a child is beaten by someone at home, how often a child's siblings are beaten, and how often a child hits his/her siblings; (d) *emotional bond* (1 = strong to 5 = weak), assessed by "Do you discuss problems with your father?" "Do you discuss problems with your mother?" and "Do you get compliments when you do something well?"; *importance of school* (1 = yes to 5 = no), assessed by "Do your parents think it is important that you go to school?" and "Do you think it is important that you learn a trade?"; (e) *general evaluation of family* (1 = good to 4.2 = bad), assessed by "How much do you like it at home?" "Do you usually deserve punishments?" "Do you have enough pocket money?" "If you have children later, will you be as stern with them as your parents are with you?" and "Do you think your parents are very stern with you?"

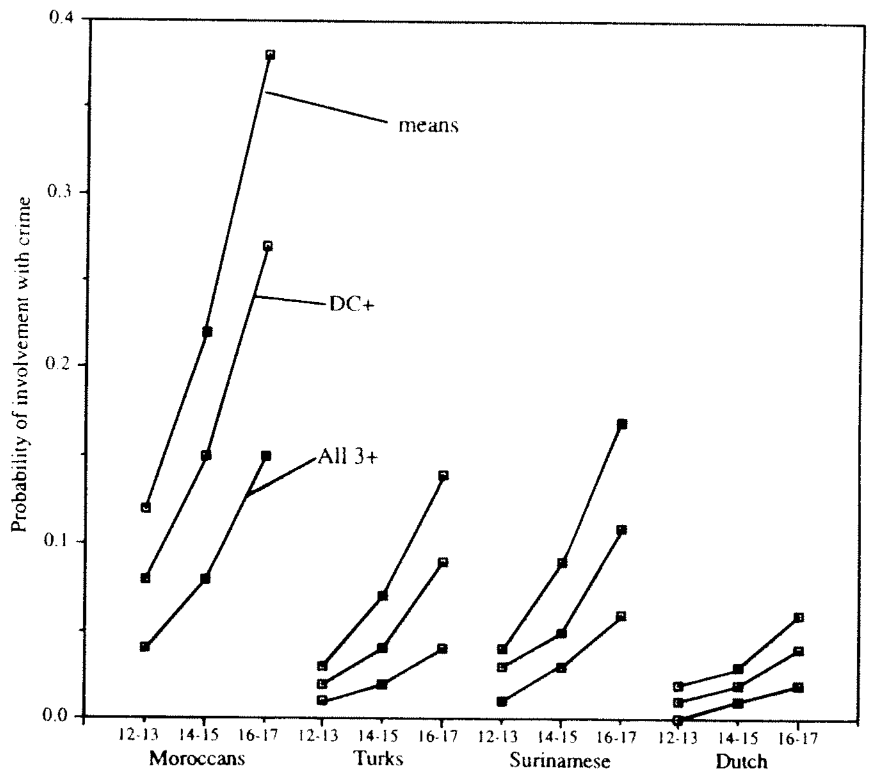

FIGURE 1. *Graphical representation of the probabilities in Table 7*

Note. For each ethnic group, the top line represents the estimates shown in Column A of Table 7, the middle line represents the estimates in Column B, and the bottom line represents the estimates in Column E

References

Bock, R. D. (1975). *Multivariate statistical methods in behavioral research.* New York: McGraw-Hill.

Clogg, C. C. (1995). Latent class models: Recent developments and prospects for the future. In G. Arminger, C. C. Clogg, & M. E. Sobel (Eds.), *Handbook of statistical modelling for the social and behavioral sciences* (pp. 311–352). New York: Plenum.

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

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

Dayton, C. M., & Macready, G. B. (1989). A latent class covariate model with applications to criterion referenced testing. In R. Langeheine & J. Rost (Eds.), *Latent trait and latent class models*, (pp. 129–143). New York: Plenum.

Dempster A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B, 39*, 1–38.

DeSarbo, W. S., & Wedel, M. (1994). A review of recent developments in latent class regression models. In R. Bagozzi (Ed.), *Handbook of marketing research* (pp. 352–388). Cambridge, MA: Blackwell.

Formann, A. K. (1992). Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association, 87*, 476–486.

Goodman, L. A. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika, 61*, 215–231.

Hagenaars, J. A. (1990). *Categorical longitudinal data. Log-linear, panel, trend and cohort analysis*, London: Sage.

Hagenaars, J. A. (1993). *Loglinear models with latent variables*. London: Sage.

Hirschi, T. (1969). *Causes of delinquency*. Berkeley: University of California Press.

Junger, M. (1990). *Delinquency and ethnicity*. Deventer, The Netherlands: Kluwer Law and Taxation.

Langeheine, R. (1988). New developments in latent class theory. In R. Langeheine & J. Rost (Eds.), *Latent trait and latent class models* (pp. 77–108) New York: Plenum Press.

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin.

Maddala, G. S. (1983). *Limited dependent and qualitative variables in econometrics*. New York: Cambridge University Press.

McCutcheon, A. L. (1994). Latent logit models with polytomous effects variables. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 353–372). London: Sage.

Mooijaart, A., & van der Heijden, P. G. M. (1992). The EM-algorithm for latent class analysis with constraints. *Psychometrika, 57*, 261–269.

Van der Heijden, P. G. M., Mooijaart, A., & de Leeuw, J. (1992). Constrained latent budget analysis. In C. C. Clogg (Ed.), *Sociological methodology 1992* (Vol. 22, pp. 279–320). Cambridge, MA: Blackwell.

Authors

PETER G. M. VAN DER HEIJDEN is Professor, Department of Methodology and Statistics, University of Utrecht, Postbus 80.140, 3508 TC Utrecht, The Netherlands; p.vanderheijden@fsw.ruu.nl. He specializes in psychology and psychometrics.

JOS DESSENS is Assistant Professor, Department of Methodology and Statistics, University of Utrecht, Postbus 80.140, 3508 TC Utrecht, The Netherlands; j.dessens@fsw.ruu.nl. He specializes in social stratification and quantitative methods for the analysis of categorical data.

ULF BOCKENHOLT is Associate Professor, Department of Psychology, University of Illinois at Urbana-Champaign, 603 East Daniel Street, Champaign, IL 61820; ubockenh@s.psych.uiuc.edu. He specializes in psychometrics and behavioral decision making.

Received February 24, 1995
Revision received June 1, 1995
Accepted June 19, 1995