

Chapter 35

Concomitant Variables in Latent Change Models

Ulf Böckenholt

University of Illinois at Urbana-Champaign

1. Introduction

A basic tenet in modeling preference behavior is that individuals differ in the ways they perceive and evaluate choice options. Latent-class analysis provides a parsimonious and flexible approach to represent these taste differences. This method decomposes a heterogeneous population of decision-makers into several homogeneous classes or subpopulations. Each decision-maker is assigned to one of the latent classes such that preference differences among members of different classes are maximized (Böckenholt, 1993; Croon & Lijckx, 1993; DeSarbo, Ramaswamy, & Lenk, 1993; Winsberg & DeSoete, 1993). The accuracy of this assignment can be greatly improved by taking into account person-specific collateral information (e.g., demographics). The consideration of collateral information is of particular interest in strategic marketing studies. If the segmentation results obtained in a latent class analysis can be characterized by demographic variables, marketing tasks such as positioning and targeting of consumer groups are much simplified (Dillon, Kumar, & Smith de Borrero, 1993; Gupta & Chintagunta, 1994).

The advantages of using collateral information become even more obvious in an analysis of longitudinal choice data. Clearly, many decisions are faced not once but repeatedly, for example, the choice of a residential location, the selection of a travel mode for a trip to work, or the purchase of consumer brands within a product-class. In these types of recurrent choice situations, individuals may not only differ in their preferences but also in the way they change their preferences over time. According to the approach by Böckenholt and Langeheine (1996), these two sources of heterogeneity can be accounted for by estimating different latent classes for each time period and allowing for switches among these classes from one period to the next. Since each latent class corresponds to a distinct preference state, preference changes over time are represented by transitions of members from one latent class to another. These latent switches can be either unrestricted, or, in a confirmatory analysis, they can be constrained to follow some hypothesized change mechanism. For example, one hypothesis of interest is that decision-makers switch among latent classes according to a Markovian process that describes the decision-makers' learning experiences with the choice options (Langeheine & van de Pol, 1988; Poulsen, 1982).

Recent applications of latent-class models to cross-sectional data demonstrate that collateral information can both improve significantly the prediction of class-membership and simplify the interpretation of a latent-class solution (Dayton & Mitchell, 1988; Hagenaars, 1990; van der Heijden, Mooijaart, & DeLeeuw, 1992; van der Heijden, Dessens, & Böckenholt, 1996). Collateral information is even more important in latent-class models that allow for changes in class-membership over time. Typically, longitudinal data are sparse

which strongly limits the precision with which individual-level changes in preference states can be estimated. By using collateral information we can obtain more accurate predictions of class-membership changes and more precisely distinguish between those decision-makers who are stable in their preferences over time and those who change their preferences. For example, in marketing applications it is of considerable interest to discriminate between brand-switchers and brand-loyal consumers because the former group is presumably more responsive to changes in marketing mix variables.

This chapter illustrates the use of concomitant variables in the analysis of longitudinal choice data. The methodological approach is based on the work by Clogg and Goodman (1984) who considered exclusively categorical concomitant variables and Dayton and Mitchell (1988) who allowed concomitant variables to be continuous. Van de Pol and Langeheine (1990) followed the approach by Clogg and Goodman (1984) when investigating the relationship between a categorical concomitant variable and the latent Markov change mechanism. This chapter extends these approaches by allowing for both continuous and discrete concomitant variables. The latent change mechanisms are restricted to fall within the class of log-linear models because this framework is general and simplifies imposing parameter constraints (Formann, 1992; van der Heijden, et al., 1996).

The remainder of this chapter is structured as follows: The latent-change Poisson model proposed by Böckenholt and Langeheine (1996) is reviewed as a framework for analyzing recurrent choice data. A logistic approach is then presented for incorporating concomitant variables in the latent change model. Finally, a consumer panel study is analyzed to illustrate the use of concomitant variables in testing hypotheses about class-membership changes in latent-class models.

2. A concomitant-variable, latent-class Poisson model

This section discusses how to relate concomitant variables to the discrete latent variables estimated by a latent-class model. This discussion is general and not limited to a particular choice model. However, because the application is concerned with recurrent choice data, we present first the latent-class Poisson model used in the application section. Suppose the data are arranged in a $(N \times T)$ frequency table with x_{vt} denoting the total number of choices observed for person v in time period t , $v = 1, \dots, N$; $t = 1, \dots, T$. To simplify the notation, it is assumed (unless otherwise stated) that these time periods are of equal duration and the number of time periods, T , is equal to 3. According to the model, the counts are independent Poisson distributed conditional on the random rate parameter λ_{vt} . Non-stationarities resulting from changes in the decision-making environment can be modeled by re-parametrizing the rate parameters as a function of external variables. To account for the variability in the individual rate parameters, it is assumed that the observations are sampled from a finite mixture of Poisson distributions. Each Poisson class is characterized by its time-homogeneous rate parameter λ_g and relative class size π_{gt} ($g = 1, \dots, G$) with $\sum_{g=1}^G \pi_{gt} = 1$.

For example, if x_{vt} belongs to class g , it follows a Poisson process with

$$p(x_{vt}; \lambda_g) = \frac{e^{-\lambda_g} \lambda_g^{x_{vt}}}{x_{vt}!}, \quad x_{vt} = 0, 1, \dots$$

Note that although the choice rate parameters do not depend on the time period, the class sizes do. Thus, time-related effects are represented exclusively by changes in class membership. As a result, the probability of observing x_{v1} , x_{v2} , and x_{v3} during three consecutive time periods is given by

$$p(x_{v1}, x_{v2}, x_{v3}; \boldsymbol{\lambda}, \boldsymbol{\pi}) = \sum_{a=1}^G \sum_{b=1}^G \sum_{c=1}^G \pi_{abc} p(x_{v1}; \lambda_a) p(x_{v2}; \lambda_b) p(x_{v3}; \lambda_c), \quad (1)$$

where π_{abc} is the joint probability of being in classes a, b, and c during the first, second, and third time period, respectively, and $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_G)'$; $\boldsymbol{\pi} = ([\pi_{abc}]; a, b, c = 1, \dots, G)$. Hypotheses about time-related effects are tested by imposing some structure on π_{abc} . For example, by setting $\pi_{abc} = 0$ when $a \neq b$, $b \neq c$, or $a \neq c$, the no-change, latent-class Poisson model is obtained as a special case of (1),

$$p(x_{i1}, x_{i2}, x_{i3}; \boldsymbol{\lambda}, \boldsymbol{\pi}) = \sum_{g=1}^G \pi_g p(x_{i1}; \lambda_g) p(x_{i2}; \lambda_g) p(x_{i3}; \lambda_g).$$

The class of log-linear models provides a powerful framework for the formulation and comparison of latent change mechanisms. Informative tests about latent change and stability in class-membership are obtained by such constraints as (quasi-)symmetry, restrictive versions of quasi-symmetry, (quasi-)independence, or linear-by-linear associations (McCutcheon & Hageaars, 1996; Langeheine & van de Pol, 1990; Vermunt, 1993). As it is shown below, the log-linear framework is also well-suited to incorporate person-specific, concomitant information in a latent change analysis.

It is convenient to impose log-linear restrictions by using Bock's (1975) multinomial logit transformation. Consequently, the joint class size probabilities are re-expressed as

$$\pi_{abc} = \frac{\exp(z_{abc})}{\sum_{a^*}^G \sum_{b^*}^G \sum_{c^*}^G \exp(z_{a^* b^* c^*})},$$

where z_{abc} is an element of the $(1 \times G^T)$ vector, $\mathbf{z} = \boldsymbol{\alpha}' \mathbf{M}$, \mathbf{M} is a $(h \times G^T)$ known contrast matrix that specifies the latent change mechanism, and $\boldsymbol{\alpha}$ is a vector containing the h unknown parameters. For example, a saturated model for $T = G = 2$ is obtained when

$$\mathbf{M} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{4} & -\frac{1}{4} & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}, \quad (2)$$

and $z_{11} = \frac{1}{2} \alpha_1 + \frac{1}{2} \alpha_2 + \frac{1}{4} \alpha_3$. The main effects are given by α_1 and α_2 and the association effect is given by α_3 . By constraining α_1 to be equal to α_2 we obtain the symmetry model of the joint class-sizes with $\pi_{ab} = \pi_{ba}$ (Bock, 1975; Haberman, 1979).

Information about decision maker i on k concomitant variables (denoted by the vector \mathbf{y}_i) can be taken into account by writing,

$$\pi_{abc|\mathbf{y}_i} = \frac{\exp(z_{abc|\mathbf{y}_i})}{\sum_{a^*, b^*, c^*}^G \exp(z_{x^* b^* c^*|\mathbf{y}_i})}$$

and the vector of conditional logits is given by

$$\mathbf{z}_i = (\mathbf{1}|\mathbf{y}_i') \mathbf{A} \mathbf{M} \quad (3)$$

where \mathbf{A} is a $([k + 1] \times h)$ matrix of unknown parameters. This parametrization assumes that the concomitant variables determine only the parameters of the latent change mechanism but that they do not affect the change mechanism itself. In some applications, it may prove necessary to relax this assumption because the change mechanism may depend on the values the concomitant variables take on. Thus, it is important to test whether the latent change mechanism applies to both the population as a whole and to the disaggregate levels of analysis defined by the concomitant variables.

Because of the typical sparseness of longitudinal data, the precision of the parameter estimates may be increased by considering more parsimonious expressions than (3). One useful approach is to use a different contrast matrix for the concomitant variables by expressing \mathbf{z}_i as

$$\mathbf{z}_i = \boldsymbol{\alpha}' \mathbf{M} + \mathbf{y}_i' \boldsymbol{\Delta} \mathbf{M}^*, \quad (4)$$

where $\boldsymbol{\Delta}$ is a $(k \times g, g \leq h)$ parameter matrix, \mathbf{M}^* is a contrast matrix that specifies the influence of the concomitant variables on the latent variables, and $\text{rank}(\frac{\mathbf{M}}{\mathbf{M}^*}) = \text{rank}(\mathbf{M})$. Consequently, (4) is a nested version of (3),

$$\mathbf{z}_i = (\mathbf{1}|\mathbf{y}_i') \begin{bmatrix} \boldsymbol{\alpha}' \\ \boldsymbol{\Delta} \mathbf{C} \end{bmatrix} \mathbf{M}, \quad (4')$$

where $\mathbf{C} = \mathbf{M}^* \mathbf{M}' (\mathbf{M} \mathbf{M}')^{-1}$, and the number of parameters is reduced by $k(h-g)$. For example, we may hypothesize that for $T = G = 2$, only the size of the association varies as a function of the external variables. In this case, \mathbf{M} is given by (2), \mathbf{M}^* by

$$\mathbf{M}^* = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 4 & 4 & 4 & 4 \end{bmatrix},$$

and

$$\mathbf{C} = [0 \ 0 \ 1].$$

3. An Application

To illustrate the concomitant-variable, latent-class Poisson model, a consumer panel study is analyzed. The data consist of the monthly numbers of purchase incidences of two brands of canned tuna-fish (Starkist in Water and Chicken of the Sea in Water in the 6.5 oz., chunk-light category) by 512 households who participated in this study for three consecutive one-month time periods. The market share of both brands is about 68%. The goal of the following investigations is to model the individual differences in purchase behavior over time and to relate these differences to the available demographic information about the panel members. Aggregate-level results were already obtained by Böckenholt and Langeheine (1996; hereafter BL) who previously modeled this data set. These results are briefly reviewed before we report the main analyses that take into account concomitant variables.

Individual differences in purchase behavior: Fitting the latent-class Poisson model to the purchase incidence data, BL found that individual variations in buying behavior during each time period are well-described by a three-class Poisson model with time-homogeneous rate parameters. However, panel members did not stay in the same class over time; instead they shifts among classes followed a complete symmetry model defined by $\pi_{abc} = \pi_{acb} = \pi_{bac} = \pi_{bca} = \pi_{cab} = \pi_{cba}$ for all a, b, c. Under this representation, an equal number of decision-makers switches from class a to class b as does from class b to a. Because the order of the time periods is irrelevant for predicting changes in class membership, switches among classes are time-homogeneous.

Table 1 contains the rate parameters estimated for the three-class symmetry model, and the time-homogeneous marginal class sizes and transition probabilities. The three classes may be labeled as heavy ($\hat{\lambda}_1 = 2.17$), medium ($\hat{\lambda}_2 = .65$), and light buyers ($\hat{\lambda}_3 = .10$). Although the size of the first and second class differ strongly ($\hat{\pi}_1 = .07, \hat{\pi}_2 = .62$), their class-membership is relatively stable ($\hat{\pi}_{11} = \hat{\pi}_{22} = .83$). According to the estimated joint class-size probabilities, $\hat{\pi}_{abc}$, about 36% of the consumers change their class membership at least once. The great majority (32%) switches only between the second and third class (i.e., from medium to light buyers) and the remaining 4% switch only between the first and third class (i.e., from heavy to light buyers). Thus, there are no transitions between the classes of heavy and medium buyers.

LC	$\hat{\lambda}_s$	$\hat{\pi}_s$	$\hat{\pi}_{a b}$		
			1	2	3
1	2.17	.07	.83	.00	.17
2	.65	.62	.00	.83	.17
3	.10	.10	.04	.35	.61

Table 1: Parameter estimates of three-class symmetry model

Effects of Covariates: Although these results provide a succinct summary of the purchase incidence data on the aggregate level, it is clear that additional information about the panel members, for example, in form of demographic variables, may prove useful for explaining the individual differences in purchase behavior. We consider four demographic variables: Number of cats owned, family size, and two dummy variables indicating whether the head of

the household is retired and whether s/he has a college degree. Table 2 contains the frequency distributions of these variables.

(a) No. of cats						(b) Family size						(c) College degree		(d) Retired	
0	1	2	3	4	≥ 5	1	2	3	4	5	≥ 6	No	Yes	No	Yes
326	124	35	19	6	2	45	162	109	128	49	19	361	151	425	87

Table 2: Frequency distribution of independent variables

Because the data table is sparse, a latent-symmetry model cannot be estimated separately for most of the combinations of the independent variables. Thus, to determine whether knowledge of the concomitant variables is helpful in assigning a consumer to one of the three classes, it was hypothesized that the four variables affect only the three marginal class sizes during each time period. This hypothesis was implemented by specifying M^* in (4) to contain linear and quadratic main effects contrasts. In addition, the corresponding parameters were constrained to be equal for the three periods to satisfy the complete symmetry constraint. To reduce the effects of extreme values of the concomitant variables, the maximum numbers for the family size and the number of cats variables were set equal to 6 and 5, respectively.

	Main effects			
	Linear		Quadratic	
No. of cats	.08	(.04)	-.04	(.03)
Family size	.01	(.04)	.01	(.02)
College	.21	(.09)	.01	(.05)
Retired	.05	(.11)	-.00	(.06)

Note: Standard errors are in parentheses.

Table 3: Parameter estimates of concomitant variables for linear and quadratic main effects

Compared to the latent change model without covariates, the drop in the log-likelihood function is not significant ($-2 (1,230.3 - 1,223.6) = 13.28, df = 8$). However, this comparison lacks power; and the standard errors of the parameter estimates in Table 3 show that under the linear main-effect, two of the four variables (no. of cats, college degree) seem to have a significant impact on the prediction of the class sizes. The value of the maximized log-likelihood function with these two variables and a linear main-effect is -1,225.4. Thus, changes in the marginal class size distributions are captured by a contrast between the two classes consisting of heavy and light buyers in the sample.

Table 4 displays the joint class sizes for two adjacent time periods as a function of the number of cats (from 0 to 2) and college degree. The rate parameters for the three classes are not included in this table because they are virtually the same as in Table 1.

According to Table 4, panel members with a college degree show a higher purchase incidence than panel members without a degree. Although there is a positive relationship between the number of cats owned and the class sizes of the heavy and light buyers, these changes in the probabilities modeled by a function that is linear in the number of cats are not well determined, as can be noted from the class sizes' standard errors (not given in Table 4). Because only a small percentage of the sample owns more than one cat, there is too little

information reliably estimate the linear effect to this group. Consequently, an alternative model in which cat ownership is treated as a dummy variable gives virtually the same fit.

<i>College = No</i>									
No. Cats	0			1			2		
LC	1	2	3	1	2	3	1	2	3
1	.04	.00	.01	.05	.00	.01	.08	.00	.01
2	.00	.48	.13	.00	.51	.12	.00	.54	.11
3	.01	.13	.21	.01	.12	.17	.01	.11	.14
<i>College = Yes</i>									
No. Cats	0			1			2		
LC	1	2	3	1	2	3	1	2	3
1	.08	.00	.01	.11	.00	.01	.15	.00	.02
2	.00	.54	.11	.00	.56	.10	.00	.57	.08
3	.01	.11	.14	.01	.10	.11	.02	.08	.08

Table 4: Joint class sizes estimates as a function of concomitant variables

The estimated joint class-membership sizes, $\hat{\pi}_{abcy}$, indicate that the probability of membership changes is highest (41%) for panel members who do not own a cat and do not have a college degree. In contrast, cat ownership in combination with the college degree reduces the probability of switching considerably. For example, according to the fitted latent-class Poisson model, only 26% of the panel members with a college degree and three cats switch class membership over time.

<i>College Degree</i>								
t_1	t_2	t_3	No (N = 361)		Yes (N = 151)		Total (N = 512)	
			obs	pred	obs	pred	obs	pred
0	0	0	35	34	26	27	32	32
0	0	1+	10	12	13	12	11	12
0	1+	0	12	12	9	12	12	12
0	1+	1+	6	8	5	9	6	8
1+	0	0	12	12	16	12	13	12
1+	0	1+	9	8	11	9	9	8
1+	1+	0	8	8	9	9	8	8
1+	1+	1+	8	8	11	11	9	9

Note: The first three columns refer to the purchase patterns for the three time periods. The grouped selection patterns are denoted by 0 (no purchase) and 1+ (at least one purchase).

Table 5: Observed and predicted purchase percentages

A final step in this analysis is to examine how well the latent symmetry constraint represents the latent change mechanism at the covariate level. Table 5 displays the (collapsed) observed and predicted purchase incidence data for panel members with and without a college degree. For example, 12% of the 361 households without a college degree buy tuna brands during the second time period only. Note that mostly non-buyers and regular buyers (i. e., consumers who purchase tuna brands during each time period) seem to differ on this variable, a finding that is well-captured by the latent-symmetry Poisson model. Table 5 also shows that the strong constraint of equal marginal distributions for the three time periods is in good agreement with the data. Further support for the latent symmetry constraint is provided by similar tests with respect to the cat ownership covariate. As with the college

degree covariate, regular buyers and non-buyers vary most on this variable: The percentage of non-buyers decreases as a function of the number of cats owned, and the opposite pattern is observed for regular buyers.

We conclude that relating covariates to latent change probabilities can yield a parsimonious and easy to interpret representation of recurrent choice data. In particular, this study demonstrated that individual differences in purchase behavior are associated with education and cat ownership. College graduates who own cats buy tuna brands more frequently and show more stability in their purchase behavior than panel members without these characteristics. In contrast, family size and retirement status were not found to be useful predictors.

4. Discussion

There are two major advantages in using collateral information in a latent class analysis of choice data. First, collateral information simplifies the interpretation of the data because both aggregate and disaggregate levels of analysis are considered. Second, additional information provided by the concomitant variables may yield more reliable parameter estimates. The latter point is especially important in the analysis of longitudinal data which are typically sparse. Extending Dayton and Macready's (1988) work on incorporating concomitant variables in latent-class models we applied Bock's (1975) multinomial logit model to allow for both individual preference differences and shifts over time. Although this approach facilitates an understanding of choice data at the disaggregate level, it is not without limitations. Most importantly, the assumption that decision-makers follow the same latent change mechanism although with different parameter values may be too restrictive in some applications. Moreover, this chapter focussed on testing hypotheses about latent change mechanisms that belong to the log-linear family. Although a large number of different hypotheses can be formulated by using this class, they are clearly not exhaustive. For example, in the reported application it may be of interest to test the competing hypothesis that the latent change probabilities follow a first-order Markov chain which is in a steady state. However, this hypothesis cannot be formulated as a log-linear model and requires therefore a different approach (see Vermunt, Langeheine, & Böckenholt, 1996).

In conclusion, because the number of joint class size probabilities is typically large and not well-determined, it is important to consider parsimonious functions of the concomitant variables as given by Model (4). This model facilitates testing the effects of discrete and continuous covariates in a straightforward fashion. As a result, the proposed modeling framework may prove well suited for identifying and understanding sources of temporal stability and change in choice behavior.

Acknowledgments

This research was partially supported by NSF grant SBR-9409531. I am grateful to Greg M. Allenby and the A. C. Nielsen Corporation for providing the data set used in this paper.

References

- Bock, R. D. (1975). *Multivariate statistical methods in behavioral research*. New York: McGraw Hill.
- Böckenholt, U. (1993). A latent class regression approach for the analysis of recurrent choice data. *British Journal of Mathematical and Statistical Psychology*, *46*, 95-118.
- Böckenholt, U. & Langeheine, R. (1996). Latent change in recurrent choice data. *Psychometrika*, *61*, 285-301.
- 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.
- Croon, M. A., & Luijkx, R. (1993). Latent structure models for ranking data. In J. Verducci & M. Fligner (eds.), *Probability models and statistical analyses for ranking data* (pp. 53-74). New York: Springer Verlag.
- Dayton, M., & Macready, G. B. (1988). Concomitant-variable latent class models. *Journal of the American Statistical Association*, *83*, 173-178.
- DeSarbo, W. S., Ramaswamy, V., & Lenk, P. (1993). A latent class procedure for the structural analysis of two-way compositional data. *Journal of Classification*, *10*, 159-193.
- Dillon, W. R., Kumar, A., & Smith de Borrero, M. (1993). Capturing individual differences in paired comparisons: An extended BTL model incorporating descriptor variables. *Journal of Marketing Research*, *30*, 42-51.
- Formann, A. K. (1992). Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association*, *87*, 307-318.
- Gupta, S., & Chintagunta, P. K. (1994). On using demographic variables to determine segment membership in logit mixture models. *Journal of Marketing Research*, *31*, 128-136.
- Haberman, S. J. (1979). *Analysis of Qualitative Data, Vol.2*. Academic Press: New York.
- Hagenaars, J. A. (1990). *Categorical Longitudinal Data Analysis*. London: Sage.
- Langeheine, R., & van de Pol, F. (1990). A unifying framework for Markov modeling in discrete space and discrete time. *Sociological Methods and Research*, *18*, 416-441.
- McCutcheon, A. L., & Hagenaars, J. A. (1994). Latent class models for comparative social research. This volume (pp. xxx-xxx).
- Poulsen, C. S. (1982). *Latent Structure Analysis with Choice Modeling Applications*. Unpublished dissertation. University of Pennsylvania.
- Van de Pol, F. J. R., & R. Langeheine (1990). Mixed Markov latent class models. In C. C. Clogg (ed.), *Sociological Methodology* (pp. 213-247). Oxford: Blackwell.
- Van der Heijden, P. G. M., Dessens, J., Böckenholt, U. (1996). Estimating the concomitant variable latent class model with the EM algorithm. *Journal of Educational and Behavioral Statistics*, *21*, 215-229.
- Van der Heijden, P. G. M., Mooijaart, & A., DeLeeuw, J. (1992). Constrained latent budget analysis. In C. C. Clogg (ed.), *Sociological Methodology* (pp. 279-320). Oxford: Blackwell.
- Vermunt, J. K. (1993). *lem: Log-linear and event history analysis with missing data using the EM algorithm*. Working paper. Tilburg University, Netherlands.
- Vermunt, J. K., Langeheine, R., & Böckenholt, U. (1996). *Discrete-time, discrete-state latent Markov models with time-constant and time-varying covariates*. Working paper. Tilburg University, Netherlands.
- Wiggins, L. M. (1973). *Panel analysis: Latent Probability Models for Attitude and Behavior Processes*. Amsterdam: Elsevier.
- Winsberg, S., & DeSoete, G. (1993). A latent class approach to fitting the weighted Euclidean model, CLASCAL. *Psychometrika*, *58*, 315-330.