PSYCHOMETRIKA—VOL. 57, NO. 2, 261–269
JUNE 1992

THE EM ALGORITHM FOR LATENT CLASS ANALYSIS WITH EQUALITY CONSTRAINTS

AB MOOIJAART

UNIVERSITY OF LEIDEN

PETER G. M. VAN DER HEIJDEN

DEPARTMENT OF EMPIRICAL AND THEORETICAL SOCIOLOGY
UNIVERSITY OF UTRECHT

The EM algorithm is a popular iterative method for estimating parameters in the latent class model where at each step the unknown parameters can be estimated simply as weighted sums of some latent proportions. The algorithm may also be used when some parameters are constrained to equal given constants or each other. It is shown that in the general case with equality constraints, the EM algorithm is not simple to apply because a nonlinear equation has to be solved. This problem arises, mainly, when equality constraints are defined over probabilities in *different* combinations of variables and latent classes. A simple condition is given in which, although probabilities in different variable-latent class combinations are constrained to be equal, the EM algorithm is still simple to apply.

Key words: latent class analysis, EM algorithm, equality constraints.

Introduction

Latent class analysis (LCA) is usually carried out by one of two methods. The first is the Newton-Raphson approach as used in Haberman's computer program LAT (Haberman, 1979), or NEWTON (Haberman, 1988). In Haberman's approach, LCA is conceived of as a specific loglinear model describing the relations between some manifest variables and a latent variable. This method requires a relatively small number of iterations, allows for the implementation of various types of constraints on the parameters, and finds asymptotic covariances for the estimates as a by-product. The second method is iterative proportional scaling (Goodman, 1974), which Dempster, Laird, and Rubin (1977) noted was a special case of the "EM algorithm" for finding maximum likelihood estimates for missing data. This strategy is used in the program MLLSA (Clogg, 1977; Eliason, 1988), LCAG (Hagenaars & Luijkx, 1990), and PANMARK (van de Pol, Langeheine, & De Jong, 1989), and although it requires many iterations to converge, each iteration is computationally very cheap and the algorithm is easily programmed. Convergence is ensured (although possibly to a nonglobal optimum), and the parameter estimates are always in the interval 0-1 (see Formann, 1978; Goodman, 1979). The question as to which approach is better is as yet unresolved (see Langeheine, 1988, for an interesting comparison).

Equality constraints and fixed-value constraints can be helpful in the interpretation phase of an analysis. Equality constraints can be used to assess whether the estimates of two or more parameters are different. Fixed-value constraints can be used to assess

The authors are grateful to the Editor and the anonymous reviewers for their helpful comments on an earlier draft of this paper. C. C. Clogg and R. Luijkx are also acknowledged for verifying our results with their computer programs MLLSA and LCAG, respectively.

Requests for reprints should be sent to Ab Mooijaart, Department of Psychology, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, THE NETHERLANDS.

0033-3123/92/0600-90238$00.75/0

whether an estimate of a parameter is significantly different from some value of theoretical interest, such as zero. Such constraints are regularly used in a second step of the analysis. In the first step, an unconstrained latent class analysis is performed; in the second, the number of parameters to be estimated is reduced using constraints.

This paper deals with the application of the EM algorithm to fit these constraints. The approach adopted will be to discuss the estimation procedures for both equality constraints and fixed-value constraints simultaneously. Thus, our results are as general as possible. A practical situation in which both fixed-value constraints and equality constraints are fitted simultaneously can be found in van de Pol and Langeheine (1990).

From the literature it appears that fixed-value and equality constraints are easily incorporated into the EM algorithm. However, we will show that this is not true in general, and in cases where probabilities of different variables and/or different classes are constrained to be equal, problems may arise. Explicitly, if equality constraints are handled incorrectly in the EM algorithm, the function value (the likelihood ratio chi-square statistic) may be increasing instead of monotonically decreasing, as it should. Examples with constructed data did indeed show that both the computer programs MMLSA and LCAG did not work properly in this respect. In a recent paper by van de Pol and Langeheine (1990), it was noticed that problems with the EM algorithm may also arise when equality constraints are imposed. They give a three-step estimation procedure for which they say: "This row-wise approach can handle almost all equality restrictions" (p. 22). This paper gives a general formulation for dealing with equality constraints in the EM algorithm.

It will be shown that for the general case with equality constraints, the EM algorithm is not simple to apply because a nonlinear equation has to be solved. It follows that for these equality constraints, the EM algorithm becomes a less attractive alternative to implement LCA, because the procedure becomes much more complicated. It seems to have gone unnoticed in the literature that for some equality constraints, nonlinear equations have to be solved in implementing the EM algorithm. Furthermore, a condition is given that, if satisfied, results in solving linear equations only, even in cases where parameters of different classes and/or variables are constrained to be equal. When this condition holds, the EM algorithm is simple to apply.

The EM Algorithm For the Unconstrained Latent Class Model

This paper begins by reviewing the estimation of the unconstrained latent class model using the EM algorithm, then discusses how it can be applied when there are constraints on the parameters.

The EM algorithm presented by Dempster, Laird, and Rubin (1977) maximizes the likelihood function, with a distinction between observed and missing data. In LCA, observed data are simply the scores on the manifest categorical variables; the missing data are the scores on a latent categorical variable whose categories are called the latent classes. Explicitly, suppose there are K latent classes indexed by k $(k = 1, \ldots, K)$, and V manifest variables indexed by $v (v = 1, \ldots, V)$, and let $\mathbf{s}$ be a vector of length V, with entries $s(v)$ denoting the category number observed for variable v assumed to take one of I_v values, $s(v) = 1, \ldots, I_v$. In LCA, the probability $\pi_{\mathbf{s}}$ of outcome vector $\mathbf{s}$ is defined as

$$\pi_{\mathbf{s}} = \sum_{k=1}^{K} \pi_{\mathbf{s},k}, \tag{1}$$

where

$$\pi_{s,k} = \pi_k \prod_{v=1}^{V} \pi_{v,s(v)|k}. \tag{2}$$

Here, π_k denotes the size of class k, and $\pi_{v,s(v)|k}$ is the probability of category $s(v)$ for variable v conditional on class k. Elements $\pi_{s,k}$ denote the unobserved probabilities of falling simultaneously in the categories denoted by vector s and the latent class k. The latent class model in (1) and (2) assumes independence between the manifest variables given level k of the latent variable.

The EM algorithm uses the likelihood for the *complete* data, that is, data for the manifest variables and the latent variable. The method consists of two steps: the expectation (E) and the maximization (M) steps. In the E-step, expectations of the unobserved complete data, conditional on the observed incomplete data matrix and the current parameter estimates, are required. In the M-step, the expected log likelihood of the complete data matrix is maximized as a function of the unknown model parameters.

Under the assumption of a multinomial distribution, the kernel of the complete-data likelihood can be written as

$$L^c = \prod_{s,k} (\pi_{s,k})^{n_{s,k}},$$

where $n_{s,k}$ denotes the number of subjects in the sample who have pattern s and fall in class k. Because class k is not observed, $n_{s,k}$ is unobserved; only the aggregate number of subjects with pattern s, n_s, is observed.

The E-step. Under the assumption of a multinomial distribution, the sufficient statistics conditional on the observed data and the model parameters are $n_{s,k}$. In the E-step the conditional expectation of $n_{s,k}$ has to be formulated. Bayes formula gives $\pi_{k|s} = \pi_k \pi_{s|k} / \sum_t \pi_t \pi_{s|t}$, where $\pi_{s|k} = \prod_v \pi_{v,s(v)|k}$ (see (2)). Consequently, since $\pi_{k|s}$ can be estimated as $n_{s,k}/n_s$, the conditional expectation of $n_{s,k}$ is $n_s \pi_{k|s} = n_s \pi_k \pi_{s|k} / (\sum_t \pi_t \pi_{s|t})$.

The M-step. In this step the complete-data log likelihood is maximized with respect to the unknown model parameters π_k and $\pi_{v,s(v)}$, with $n_{s,k}$ replaced by the conditional expectation just found in the E-step, denoted as $n^+_{s,k}$. So we maximize (see (2))

$$\log L^c = \sum_{s,k} n^+_{s,k} \log(\pi_{s,k}) = \sum_{s,k} n^+_{s,k} \log\left(\pi_k \prod_v^V \pi_{v,s(v)|k}\right)$$

$$= \sum_{s,k} n^+_{s,k} \log \pi_k + \sum_{s,k} n^+_{s,k} \sum_v^V \log \pi_{v,s(v)|k}.$$

We optimize log Lc using Lagrange multipliers.

Estimation of π_k. The side condition for π_k is $\sum_k \pi_k = 1$. The Lagrangian can be written as

$$f_1(\pi_k) = \sum_{s,k} n^+_{s,k} \log \pi_k - \alpha\left(\sum_k \pi_k - 1\right),$$

so $\partial f_1/\partial \pi_k = \sum_s (n_{s,k}^+/\pi_k) - \alpha$. From equating $\partial f_1/\partial \pi_k = 0$ and by the side condition, it follows that $\alpha = N$, where N is the sample size. Thus, π_k can be estimated as $\pi_k^+ = \sum_s n_{s,k}^+/N$.

Estimation of $\pi_{v,s(v)|k}$. It is easy to verify that

$$\sum_{s,k} n_{s,k}^+ \sum_{v}^{V} \log \pi_{v,s(v)|k} = \sum_{v}^{V} \sum_{i}^{I_v} \sum_{k}^{K} n_{v,i,k}^+ \log \pi_{v,i|k},$$

where for notational ease, the unknown conditional probabilities are now denoted as $\pi_{v,i|k}$, and for which the side condition, $\sum_i \pi_{v,i|k} = 1$, holds. For the same reason, $n_{v,i,k}$ is used instead of $n_{s,k}$, and is defined as the number of subjects in category i of variable v and, simultaneously, in class k. Obviously, $n_{v,i,k}^+$ is the updated value of $n_{v,i,k}$ from the E-step.

Thus, the function to be optimized becomes

$$f_2(\pi_{v,i|k}) = \sum_{v}^{V} \sum_{i}^{I_v} \sum_{k}^{K} n_{v,i,k}^+ \log \pi_{v,i|k} - \sum_{v}^{V} \sum_{k}^{K} \beta_{v,k}\left(\sum_{i}^{I_v} \pi_{v,i|k} - 1\right),$$

where the $\beta_{v,k}$'s are the Lagrange multipliers. Because $\partial f_2(\pi_{v,i|k})/\partial \pi_{v,i|k} = (n_{v,i,k}^+/\pi_{v,i|k}) - \beta_{v,k}$, setting this equation to zero and using the side condition gives $\beta_{v,k} = \sum_i n_{v,i,k}^+$. It follows that $\pi_{v,i|k}$ can be estimated as $\pi_{v,i|k}^+ = n_{v,i,k}^+/\sum_j n_{v,j,k}^+$.

The EM Algorithm in Constrained Latent Class Analysis

In this section we discuss the estimation of probabilities when some of the probabilities are constrained. Constraints on class sizes turn out not to be a problem, and therefore we concentrate on constraints for the conditional probabilities. Two types of constraints are considered: fixed-value constraints for one or more conditional probabilities $\pi_{v,i|k}$, and equality constraints for one or more sets of conditional probabilities. The equality constraints are of the general form $\pi_{v,i|k} = \pi_{w,j|l}$. Note that this means that the conditional probabilities of different variables and of different classes may be constrained to be equal. For notational ease, we drop the superscript '+' from the conditional expectations $n_{v,i,k}^+$ of the sufficient statistics from the E-step. For the estimation of the conditional probabilities $\pi_{v,i|k}$, the function to be maximized is

$$f(\pi_{v,i|k}) = \sum_{v,i,k}^{V,I_v,K} n_{v,i,k} \log \pi_{v,i|k}. \tag{3}$$

We introduce more notation. There are three sets of parameters: a set of fixed parameters, a set of free parameters, and a set of equal parameters. The set of elements $\pi_{v,i|k}$ constrained to fixed-values will be denoted by F, and the set of free parameters by G. In total, there are L equality sets, denoted by $E_l (l = 1, \ldots, L)$, where E_l consists of elements $\pi_{v,i|k}$ that are constrained to be equal. Each $\pi_{v,i|k}$ belongs either to one equality set or to none.The union of the sets E_l is denoted as E, and the elements $\pi_{v,i|k}$ in set E_l are equal to π_{E_l}. (It is assumed that if elements are fixed, they belong only to F, even when the fixed-values are equal.) In addition, we define subsets $F_{v,k}$ and $G_{v,k}$, of F and G, respectively, where the elements of these subsets belong to variable v and class k only. We further define for a set X: $n_X \equiv \sum n_{v,i,k}$, with $\pi_{v,i|k} \in X$, where X may be any set defined above. For instance, $n_{G_{v,k}}$ is the sum of the elements in the frequency table of variable v and class k that correspond to the free elements. Obviously, $\sum_l n_{E_l} = n_E$. We also define $n_{v,k}$ as $\sum_i n_{v,i,k}$, and $d_{l,v,k}$ as the number of elements of variable v and class k that belong to E_l.

The function $f(\pi_{v,i|k})$ in (3) is to be maximized over the unknown parameters π_{E_l} and the unknown free parameters $\pi_{v,i|k}$, so rewrite (3) as

$$f(\pi_{E_l}, \pi_{v,i|k}) = \sum_{l}^{L} n_{E_l} \log \pi_{E_l} + \sum_{\substack{v,i,k \\ v,i,k \in G}} n_{v,i,k} \log \pi_{v,i|k} + \sum_{\substack{v,i,k \\ v,i,k \in F}} n_{v,i,k} \log \pi_{v,i|k}, \tag{4}$$

and maximize (4) using Lagrange multipliers. There are VK different combinations of variables and latent classes, indexed by v and k. The restriction that holds for each variable-latent class combination is

$$\left(\sum_{l}^{L} d_{l,v,k} \pi_{E_l} + \sum_{\substack{i \\ v,i,k \in G}}^{I_v} \pi_{v,i|k} \right) = \left(1 - \sum_{\substack{i \\ v,i,k \in F}}^{I_v} \pi_{v,i|k} \right) \equiv c_{v,k}. \tag{5}$$

The constrained function including the Lagrange multipliers can now be written as

$$f^*(\pi_{E_l}, \pi_{v,i|k}) = \left(\sum_{l}^{L} n_{E_l} \log \pi_{E_l} \right) + \left(\sum_{\substack{v,i,k \\ v,i,k, \in G}} n_{v,i,k} \log \pi_{v,i|k} \right) - \sum_{v,k}^{V,K} \alpha_{v,k} \left(\sum_{l}^{L} d_{l,v,k} \pi_{E_l} + \sum_{\substack{i \\ v,i,k \in G}}^{I_v} \pi_{v,i|k} - c_{v,k} \right).$$

The derivatives of this function with respect to the unknown parameters are

$$\frac{\partial f^*(\pi_{E_l}, \pi_{v,i|k})}{\partial \pi_{E_l}} = \frac{n_{E_l}}{\pi_{E_l}} - \sum_{v,k}^{V,K} d_{l,v,k} \alpha_{v,k}, \tag{6}$$

$$\frac{\partial f^*(\pi_{E_l}, \pi_{v,i|k})}{\partial \pi_{v,i|k}} = \frac{n_{v,i,k}}{\pi_{v,i|k}} - \alpha_{v,k}, \qquad i, k \in G. \tag{7}$$

Equating the derivatives (6) and (7) to zero and solving for the parameters, gives

$$\pi_{E_l} = \frac{n_{E_l}}{\sum_{v,k}^{V,K} d_{l,v,k} \alpha_{v,k}}, \tag{8}$$

$$\pi_{v,i|k} = \frac{n_{v,i,k}}{\alpha_{v,k}}, \qquad i, k \in G. \tag{9}$$

So by solving for the VK Lagrange multipliers, $\alpha_{v,k}$, we obtain from (5) and (8) through (9),

$$\sum_{l}^{L} \frac{d_{l,v,k} n_{E_l}}{\sum_{w,t}^{V,K} d_{l,w,t} \alpha_{w,t}} + \frac{n_{G_{v,k}}}{\alpha_{v,k}} = c_{v,k}, \tag{10}$$

which defines VK equations for VK unknown parameters $\alpha_{v,k}$. Solving for these parameters from (10) and substituting into (8) and (9) gives the solutions of the unknown model parameters. So (10) is the basic formula that has to be solved. However, in the *general* case, a simple (explicit) solution for the parameters does not exist, and solving for $\alpha_{v,k}$ from (10) needs an iterative procedure such as the Newton-Raphson method. Therefore, we will distinguish four special cases.

Special Cases

In some cases, there does exist an explicit solution for $\alpha_{v,k}$ from (10), and thus, there are explicit solutions for the unknown parameters; in others, an iterative procedure has to be used for solving $\alpha_{v,k}$ from (10).

Case 1: No Equality Constraints for the Probabilities in the Variable-Latent Class Combination of Variable v in Class k

This case considers the estimation of conditional probabilities of variable v in latent class k when there may be fixed probabilities in the variable–latent class combination, but none of the probabilities in the variable–latent class combination is constrained to equal any other probabilities. This means that $d_{l,v,k} = 0$, for all l. Equality constraints for other variable–latent class combinations may hold because such restrictions do not influence the estimation procedure for the free parameters in the variable–latent class combination of variable v and class k. Solving $\alpha_{v,k}$ from (10) is simple, and the solutions for the model parameters space become

$$\pi^{+}_{v,i|k} = \frac{n_{v,i,k}}{s_{v,k}}, \tag{11}$$

where $s_{v,k} \equiv (n_{v,k} - n_{F_{v,k}})/c_{v,k}$ is a scaling factor for the variable–latent class combination of variable v and class k that guarantees that the parameters in this variable–latent class combination sum to one. Obviously, if there are no fixed probabilities, then $c_{v,k} = 1$ and $n_{F_{v,k}} = 0$, and (11) gives the same solution mentioned earlier.

Case 2. Probabilities in Different Variable–Latent Class Combinations are Not Constrained to be Equal

Note, there may be several equality sets, and within each variable–latent class combination there may be more than one equality set, but probabilities of *different* variable–latent class combinations are not constrained to be equal. Thus, if $d_{l,v,k} \neq 0$, then $d_{l,w,t} = 0$, for $(w, t) \neq (v, k)$, and $\alpha_{v,k}$ can again be solved easily from (10). Substituting $\alpha_{v,k}$ into (8) and (9) gives

$$\pi^{+}_{E_l} = \frac{n_{E_l}}{d_{l,v,k} s_{v,k}},$$

$$\pi^{+}_{v,i|k} = n_{v,i,k}/s_{v,k},$$

with $s_{v,k}$ as defined before. Note that all parameters in the variable–latent class combination of variable v and class k are estimated, whether free or constrained to be equal.

Remark 1. Because parameters in variable–latent class combinations for which Case 1 or 2 hold can be solved easily, we exclude these variable–latent class combinations from consideration and assume that equality sets are defined over different variable–latent class combinations. Thus a summation over the variable–latent class combinations in the discussion to follow should be interpreted as a summation over all variable–latent class combinations, with the variable–latent class combinations mentioned in Cases 1 or 2 excluded. For instance, $n_G \equiv \sum_{v,i,k} n_{v,i,k}$, with $\pi_{v,i|k} \in G$ and with the summation over v and k only for those variable–latent class combinations for which Cases 1 and 2 do not hold.

Remark 2 By leaving out those variable–latent class combinations that follow Cases 1 and 2, an interesting equation can be derived from (10). Multiplying the left and right hand side in (10) by $\alpha_{v,k}$, and summing over v and k, results in

$$n_E + n_G = \sum_{v}^{V} \sum_{k}^{K} \alpha_{v,k} c_{v,k}. \tag{12}$$

Case 3: Probabilities of Different Variable–Latent Class Combinations are Constrained to be Equal, and $d_{l,v,k} = d_l c_{v,k}$

This case is particularly important, because here there are equal probabilities in *different* variable–latent class combinations, although there is a simple solution for the parameters. The most common situation in which this constraint holds is when $c_{v,k} = 1$, implying there are no fixed elements, or the fixed elements are equal to zero. Then, $d_{l,v,k} = d_l$, which means that the number of elements in each variable–latent class combination that has elements in the set E_l is equal to d_l.

Substituting $d_{l,v,k} = d_l c_{v,k}$ into (10), and using (12) gives

$$\alpha_{v,k} = \frac{(n_E + n_G) n_{G_{v,k}}}{c_{v,k} n_G}.$$

Thus, from (8) and (9) we can derive explicit expressions for the parameters π_{E_l} and $\pi_{v,i|k}$ as

$$\pi^{+}_{E_l} = \frac{n_{E_l}}{d_l(n_E + n_G)},$$

$$\pi^{+}_{v,i|k} = \frac{c_{v,k} n_{v,i,k} n_G}{(n_E + n_G) n_{G_{v,k}}},$$

taking into account Remark 1, again.

Case 4: Probabilities of Different Variable–Latent Class Combinations are Constrained to be Equal.

Note that Case 3 is a special case of Case 4. In this most general case, it is not always simple to estimate the parameters. For illustrational reasons, we start with the example of $L = 1$, where it may hold that probabilities of different variable–latent class combinations are constrained to be equal. Combining (8) and (10) gives

$$d_{1,v,k}\pi_{E_1} + \frac{n_{G_{v,k}}}{\alpha_{v,k}} = c_{v,k}, \qquad v = 1, \ldots, V;\ k = 1, \ldots, K. \tag{13}$$

From (13), we can express $\alpha_{v,k}$ as a function of π_{E_1}. Combining (12) and (13) gives

$$\sum_{v}^{V}\sum_{k}^{K} \frac{c_{v,k} n_{G_{v,k}}}{c_{v,k} - d_{1,v,k}\pi_{E_1}} = n_{E_1} + n_G.$$

This is just one nonlinear equation with one unknown parameter, π_{E_1}. Of course, the parameter can be solved for, and as experience has shown, this can be done quite efficiently by the Newton–Raphson algorithm. Once π_{E_1} has been found, $\alpha_{v,k}$ can be derived from (13), and $\pi_{v,i|k}$ can be derived from (9). However, in the EM algorithm, estimating the parameters in the M-step has to be done many times, so solving a nonlinear system in each step may become a computational burden. If there is more than one equality set, the situation may be even worse since a nonlinear system of L equations has to be solved in each step.

Conclusions

In words our result is:

1. In cases where each of the equality constraints holds only for the parameters in one variable–latent class combination, the standard EM estimation procedure gives correct results.
2. In cases where the number of elements of an equality set is equal for different variable–latent class combinations, the standard EM estimation procedure is correct, assuming that the fixed elements are zero. When the fixed elements are non-zero, the condition is more complicated. See Case 3 above.
3. In all other cases, for each EM-step, estimation of the parameters has to be done by an iterative procedure.

Manuals of computer programs, like MLLSA (Clogg, 1977; Eliason, 1988) and LCAG (Hagenaars & Luijkx, 1990), unfortunately do not give enough detailed information about how the program deals with equality constraints in general. In many practical situations, one will be interested in the classes of constraints defined by Cases 1, 2, and 3 above. In these instances the EM algorithm can still be applied in a straightforward way. However, in the general case, equality constraints cannot be incorporated into the EM algorithm in a straightforward manner since the unknown parameters cannot be estimated as simple weighted sums of latent proportions. An iterative procedure has to be used in each M-step of the EM algorithm, and because the EM algorithm often needs many iterations to converge, this can make the algorithm even slower than it normally is. The contrary holds for the estimation of LCA by the Newton-Raphson algorithm since computation time generally decreases if the number of parameters to be estimated decreases. For Cases 1, 2, and 3, the EM algorithm seems preferable because it is computationally cheap in each iteration step and easy to program. For the general case of equality constraints in LCA, more research is needed to decide whether a modified EM algorithm or another method (like the Newton-Raphson algorithm) is to be prefered.

References

Clogg, C. C. (1977). *Unrestricted and restricted maximum likelihood latent structure analysis: A manual for users* (Working Paper No. 1977-09) University Park, PA: Pennsylvania State University.

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.

Eliason, S. R. (1988). *The categorical data analysis system. Version 3.00A user's manual.* University Park, PA: Pennsylvania State University, Department of Sociology.

Formann, A. K. (1978). A note on parameter estimation for Lazarsfeld's latent class analysis. *Psychometrika, 43*, 123–126.

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

Goodman, L. A. (1979). A note on the estimation of parameters in latent structure analysis. *Psychometrika, 44*, 123–128.

Haberman, S. J. (1979). *Analysis of qualitative data* (2 vols.). New York: Academic Press.

Haberman, S. J. (1988). A stabilized Newton-Raphson algorithm for loglinear models for frequency tables derived by indirect observation. In C. C. Clogg (Ed.), *Sociological methodology*, 1988 (pp. 193–211). Washington, DC: American Sociological Association.

Hagenaars, J. A., & Luijkx, R. (1990). *LCAG* (Working Paper Series #17). Tilburg: Tilburg University, Department of Sociology.

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.

van de Pol, F. J. R., & Langeheine, R. (1990). Mixed markov latent class models. In C. C. Clogg (Ed.), *Sociological methodology, 1990* (pp. 213–247). Oxford: American Sociological Association.

van de Pol, F. J. R., Langeheine, R., & De Jong, W. (1989). *PANMARK user manual, Panel analysis using Markov chains, version 1.5.* Voorburg: Netherlands Central Bureau of Statistics.

Manuscript received 5/29/90
Final version received 7/15/91