9

CONSTRAINED LATENT BUDGET ANALYSIS

*Peter G. M. van der Heijden**
Ab Mooijaart†
Jan de Leeuw‡

A budget is defined as a row of a two-way table consisting of conditional probabilities adding up to one. The latent budget model approximates the observed budgets of a table by a lower number of underlying, or latent, budgets. The model was originally proposed by Clogg (1981) in the context of square social mobility tables. In this paper we discuss the model in the context of two-way contingency tables. We extend the latent budget model by imposing constraints upon the parameters. Special attention is given to imposing multinomial logit constraints on the latent budget parameters. We show what these constraints imply for the interpretation of the latent budget model as a loglinear model for the latent probabilities. We discuss two examples.

1. INTRODUCTION

In a two-way contingency table, the row and the column variables regularly play different roles: One can be considered an ex-

We gratefully acknowledge the comments of Peter Marsden and the reviewers on an earlier version of the paper. We also thank Ad Kerkhof, who helped us interpret the suicide example, and Murray Pearson, who corrected the English.

*University of Utrecht, The Netherlands
†University of Leiden, The Netherlands
‡University of California, Los Angeles

planatory variable and the other a response variable. In such a situation we are interested in the dependence of the response variable on the explanatory variable. One way to study this asymmetric relation between the two variables is by comparing the proportions conditional on each level of the explanatory variable.

In this paper we examine a model for the conditional probabilities, namely, the latent budget model (see van der Heijden, Mooijaart, and de Leeuw 1989; de Leeuw, van der Heijden, and Verboon 1990). A vector with conditional proportions that add up to one is called an *observed* budget; for each row there is an observed budget that specifies the observed proportions of the response variable for that row. The latent budget model approximates the observed budgets by a mixture of one or more unknown or *latent* budgets. There are two types of parameters in the latent budgets: (a) conditional probabilities that add up to one (probabilities which have to be estimated) and (b) mixing parameters that define how the latent budgets are mixed to approximate as closely as possible the observed budgets.

In the context of square social mobility tables, Clogg (1981) presented the latent budget model as a reparameterization of latent class analysis. Unaware of this earlier work de Leeuw and van der Heijden (1988) independently found the latent budget decomposition for the analysis of so-called time-budget data. Time-budget data are a specific type of constant row-sum data (also called compositional data) in which the matrix to be analyzed comprises (groups of) individuals in the rows, activities in the columns, and proportions of time spent by individuals on the activities in the cells. In the same context of time-budget analysis, de Leeuw et al. (1990) discussed the identifiability of the model in more detail and demonstrated the relation of latent budget analysis to logcontrast principal component analysis (Aitchison 1986), which is another method for the analysis of compositional data.

Preliminary results of latent budget analysis in the general context of contingency tables can be found in van der Heijden et al. (1989). De Leeuw and van der Heijden (1991) discussed the relationship between latent budget analysis, (simultaneous) latent class analysis, and (a maximum likelihood version of) correspondence analysis (Goodman 1985, 1986; Gilula and Haberman 1986, 1988). Van der Heijden (1991) discussed the relationship between versions of the latent budget model, the RC(M)-association model, and correspondence analysis, which deal with structural zero cells in two-way tables.

Van der Heijden et al. (1989) also showed how latent budget analysis can be used to analyze higher-way contingency tables. To analyze higher-way tables, one must subdivide the variables into two subgroups: explanatory variables and response variables. The variables in each subgroup are treated as a joint variable.

In this paper we extend the latent budget model by imposing constraints upon the parameters of the model. We consider fixed-value constraints, equality constraints, and linear-logistic constraints. Constraining parameters in models is important for several reasons. First, imposing further constraints (i.e., if they are legitimate) often simplifies the model because it reduces the number of parameters to be interpreted. Second, substantive research questions can sometimes be formulated as constraints on the latent budget model. Testing the admissibility of these constraints provides answers to the research questions. Third, constraining parameters reduces the standard errors of the unconstrained parameter estimates.

In section 2 we define the latent budget model and discuss its relevant properties. In section 3 we discuss the three types of constraints. Then we discuss the identifiability of the model when constraints are used and the degrees of freedom. In the examples, we give special attention to the analysis of higher-way tables, where by imposing constraints, we can use the factorial structure of joint variables.

2. UNCONSTRAINED LATENT BUDGET MODEL

2.1. *Introduction*

To present the latent budget model formally, we introduce some notation. Let observed proportions be denoted as p_{ij}, where i ($i = 1, \ldots, I$) indexes the levels of the explanatory (row) variable A, and j ($j = 1, \ldots, J$) indexes the levels of the response (column) variable B. If we sum over an index, we will replace the index by a plus sign: $p_{i+} = \Sigma_j p_{ij}$. The proportions p_{ij} are derived from frequencies n_{ij} as $p_{ij} = n_{ij}/N$, where $N = n_{++}$.

The conditional proportions we are interested in are p_{ij}/p_{i+}. The independence model is often used as a baseline in the study of the dependence of B on A. Independence implies for the theoretical probabilities π_{ij} that $\pi_{ij}/\pi_{i+} = \pi_{+j}$, which shows that under independence, the conditional probability for column j is π_{+j} irrespective of the level of the explanatory variable. Thus, the *dependence* of the response variable on the explanatory variable can be studied by comparing values p_{ij}/p_{i+} for different i. The logit model can be used to study the dependence of the response variable B on the explanatory variable A if the response variable is dichotomous. The multinomial logit model can be used if the response variable is polytomous (see, for example, Bock 1975; Haberman 1979; Agresti 1990). The independence model is a special case of such models. The independence model is also a special case of the latent budget model.

In an observed budget, with the conditional proportions $(p_{i1}/p_{i+}, \ldots, p_{ij}/p_{i+}, \ldots, p_{iJ}/p_{i+})$, row budget i is the conditional distribution of the column categories for row i. The latent budget model describes the *theoretical* budgets with elements π_{ij}/π_{i+} as a mixture of T *latent* budgets, indexed by t ($t = 1, \ldots, T$). Let the latent budget t have parameters $\pi_{jt}^{\bar{B}X}$, where the bar over B shows that these latent budget parameters are conditional probabilities interpreted as follows: The parameter $\pi_{jt}^{\bar{B}X}$ specifies the probability of level j for response variable B given latent budget t. Let the mixture be defined by mixture parameters $\pi_{it}^{A\bar{X}}$, where the bar over variable X indicates that the mixture parameters are conditional probabilities interpreted as follows: The parameter $\pi_{it}^{A\bar{X}}$ specifies the probability that an observation falls into latent budget t given level i for the explanatory variable A. The model is defined thus:

$$\frac{\pi_{ij}}{\pi_{i+}} = \sum_{t=1}^{T} \pi_{it}^{A\bar{X}} \pi_{jt}^{\bar{B}X}, \tag{1}$$

with constraints

$$\begin{aligned} &\sum_{t=1}^{T} \pi_{it}^{A\bar{X}} = 1 \quad \text{for } i = 1, \ldots, I, \pi_{it}^{A\bar{X}} \geq 0, \\ &\sum_{j=1}^{J} \pi_{jt}^{\bar{B}X} = 1 \quad \text{for } t = 1, \ldots, T, \pi_{jt}^{\bar{B}X} \geq 0. \end{aligned} \tag{2}$$

If the number of latent budgets T is 1, then (1) is equivalent to the independence model: In this case, $\pi_{it}^{A\bar{X}} = 1$ for all i and $\pi_{jt}^{\bar{B}X} = \pi_{+j}$. If T

= min (I,J), then the model is equivalent to a saturated model: In this case it does not impose constraints.

On the assumption that the observations are distributed multinomially for each level of the explanatory variable A, maximum likelihood estimates can be derived (see below). The model can be tested against the unconstrained alternative using the Pearson chi-square test and the likelihood ratio chi-square test. If the model is true, then the test statistic follows asymptotically a chi-squared distribution with $(I - T)(J - T)$ degrees of freedom. The conditional test of the model with $T = n$ latent budgets, given that the model with $T = n + m$ latent budgets is true ($n,m \geq 1$, $n + m < \min(I,J)$), does *not* follow asymptotically a chi-squared distribution because we are working in the domain of mixture models. (See Aitkin, Anderson, and Hinde [1981] and Everitt [1988] for discussions of this problem in latent class analysis.)

2.2. *Example*

To motivate the model, we analyze German suicide data. Van der Heijden and de Leeuw (1985) used correspondence analysis to analyze these data. The row variable is a cross-classification of two explanatory variables—namely, age group and sex—and the column variable, or response variable, is cause of death. The data are given in Table 1. They were collected by the German Office for Statistics in Western Germany for the years 1974 to 1977, and are provided by Heuer (1979, Table 1). The latent budget model aims to find the latent budgets of cause-of-death categories that have generated the $2 \times 17 = 34$ observed budgets of cause-of-death categories. The latent budgets can be interpreted as typical cause-of-death distributions.

For the model with one latent budget, $G^2 = 10{,}332.9$ (df = 264). For two latent budgets, $G^2 = 4{,}595.4$, $df = 224$; for three, $G^2 = 1{,}085.9$, $df = 186$; and for four, $G^2 = 465.7$, $df = 150$. We use the test statistics only as descriptive measures for three reasons. First, the sample size is large ($n = 53{,}210$). Second, we are analyzing population data, so we don't have to make inferences from the sample to the population. Third, the observations are most likely not completely independent, since it is known that suicides generate new suicides with similar characteristics (here, suicides of the same age and sex and the same cause of death). We use the model with one

TABLE 1
Suicide Behavior: Age by Sex by Cause of Death

Age	Cause of Death[a] 1	2	3	4	5	6	7	8	9	Total
Males										
10–15	4	0	0	247	1	17	1	6	9	285
15–20	348	7	67	578	22	179	11	74	175	1,461
20–25	808	32	229	699	44	316	35	109	289	2,561
25–30	789	26	243	648	52	268	38	109	226	2,399
30–35	916	17	257	825	74	291	52	123	281	2,836
35–40	1,118	27	313	1,278	87	293	49	134	268	3,567
40–45	926	13	250	1,273	89	299	53	78	198	3,179
45–50	855	9	203	1,381	71	347	68	103	190	3,227
50–55	684	14	136	1,282	87	229	62	63	146	2,703
55–60	502	6	77	972	49	151	46	66	77	1,946
60–65	516	5	74	1,249	83	162	52	92	122	2,355
65–70	513	8	31	1,360	75	164	56	115	95	2,417
70–75	425	5	21	1,268	90	121	44	119	82	2,175
75–80	266	4	9	866	63	78	30	79	34	1,429
80–85	159	2	2	479	39	18	18	46	19	782
85–90	70	1	0	259	16	10	9	18	10	393
90+	18	0	1	76	4	2	4	6	2	113
Females										
10–15	28	0	3	20	0	1	0	10	6	68
15–20	353	2	11	81	6	15	2	43	47	560
20–25	540	4	20	111	24	9	9	78	67	862
25–30	454	6	27	125	33	26	7	86	75	839
30–35	530	2	29	178	42	14	20	92	78	985
35–40	688	5	44	272	64	24	14	98	110	1,319
40–45	566	4	24	343	76	18	22	103	86	1,242
45–50	716	6	24	447	94	13	21	95	88	1,504
50–55	942	7	26	691	184	21	37	129	131	2,168
55–60	723	3	14	527	163	14	30	92	92	1,658
60–65	820	8	8	702	245	11	35	140	114	2,083
65–70	740	8	4	785	271	4	38	156	90	2,096
70–75	624	6	4	610	244	1	27	129	46	1,691
75–80	495	8	1	420	161	2	29	129	35	1,279
80–85	292	3	2	223	78	0	10	84	23	715
85–90	113	4	0	83	14	0	6	34	2	256
90+	24	1	0	19	4	0	2	7	0	57
Total	17,565	253	2,154	20,377	2,649	3,118	937	2,845	3,313	53,211

[a]Cause-of-death categories: 1 = ingestion of solid or liquid matter; 2 = gas poisoning at home; 3 = poisoning by other gas; 4 = hanging, strangling, suffocation; 5 = drowning; 6 = guns or explosives; 7 = knives, etc.; 8 = jumping; 9 = other methods.

latent budget as the baseline model, which assumes that age-sex combinations are independent of the cause of death. Models with more than one latent budget describe the dependence of the response variable on the explanatory variables. The model with two latent budgets fits .555 of this dependence ([10,332.9 − 4,595.4] ÷ 10,332.9), the model with three latent budgets fits .895, and the model with four latent budgets fits .955 of this dependence. We focus on the solution with three latent budgets, keeping in mind that this solution shows only the main aspects of dependence. Including a fourth latent budget leads to a further gain of only 6 percent. The parameter estimates are given in Tables 2a and 2b.

We first study the latent budgets that have generated the expected budgets. The latent budget parameters $\hat{\pi}_{jt}^{\bar{B}X}$ can be interpreted cursorily by comparing them with their corresponding marginal probabilities $\hat{\pi}_{+j}$. If $\hat{\pi}_{jt}^{\bar{B}X} > \hat{\pi}_{+j}$, then latent budget t is characterized by category j (among others) in the sense that, given t, a much higher probability of j is observed than if we had no information about t. This shows that the first latent budget is characterized by relatively high conditional probabilities of suicide by ingestion of solid or liquid matter, including medicine ($\hat{\pi}_{jt}^{\bar{B}X} = .530$ versus $\hat{\pi}_{+j} = .330$), by gas poisoning (gas home, .011 versus .005; gas other, .135 versus .040), by guns and explosives (.104 versus .059), and by other methods (.141 versus .062). This latent budget is used relatively more often by younger adults: For example, for males aged 15–20, $\hat{\pi}_{it}^{A\bar{X}} = .519$; for males aged 20–25, $\hat{\pi}_{it}^{A\bar{X}} = .666$; for females aged 15–20, $\hat{\pi}_{it}^{A\bar{X}} = .458$, and so on).

The second latent budget, used mainly by females (see estimates $\hat{\pi}_{it}^{A\bar{X}}$ in second column for females), gives estimates relatively higher than marginal probabilities for drowning (.099 versus .050), ingestion of solid or liquid matter (.437 versus .330), and jumping (.084 versus .053), methods that are relatively less violent. The third latent budget is used mainly by males and gives relatively higher estimates for hanging (.840 versus .383), guns and explosives (.090 versus .059), knives (.029 versus .018), and so on.

To understand how the expected budgets are constructed from the latent budgets, we have to consider the row parameters. A cursory overview of all the row parameter estimates $\hat{\pi}_{it}^{A\bar{X}}$ can be obtained by comparing them with the parameter estimates $\hat{\pi}_{t}^{X}$: .296, .396, and .308. These are the probabilities of each of the latent

TABLE 2a
Latent Budget Analysis Parameter Estimates for Data in Table 1
(Row Parameters $\hat{\pi}_{it}^{AX}$)

Age	Males $t = 1$	Males $t = 2$	Males $t = 3$	Females $t = 1$	Females $t = 2$	Females $t = 3$
10–15	.026	.000	.974	.278	.642	.081
15–20	.519	.000	.481	.458	.542	.000
20–25	.666	.000	.334	.408	.592	.000
25–30	.671	.000	.329	.428	.572	.000
30–35	.626	.031	.343	.300	.700	.000
35–40	.543	.060	.396	.303	.697	.000
40–45	.491	.054	.455	.154	.834	.012
45–50	.460	.033	.507	.109	.882	.009
50–55	.354	.128	.518	.080	.878	.042
55–60	.290	.187	.522	.056	.909	.036
60–65	.228	.229	.544	.016	.937	.047
65–70	.126	.318	.556	.000	.931	.069
70–75	.074	.374	.552	.000	.967	.033
75–80	.037	.391	.573	.000	.978	.022
80–85	.006	.482	.511	.000	1.000	.000
85–90	.000	.414	.586	.000	1.000	.000
90+	.029	.346	.625	.000	.999	.001
$\hat{\pi}_t^X$	.296	.396	.308			

budgets when there is no information about the level of the row variable. The parameter estimates $\hat{\pi}_t^X$ are the weighted averages of the row parameters $\hat{\pi}_{it}^{AX}$. Comparison of $\hat{\pi}_{it}^{AX}$ and $\hat{\pi}_t^X$ shows that the first latent budget is used predominantly by males aged 15 to 55 and by females aged 15 to 40. It is hardly used by young boys (aged 10–15) by males older than 70, or by women older than 40. The second latent budget is used mainly by males aged 80 to 90 and by females of all ages, but it is hardly used by males under 40. The third latent budget is predominantly used by males of all ages and is hardly used by females. Roughly speaking, then, the first latent budget is used mainly by younger adults, the second is used mainly by females and older males, and the third is used almost exclusively by males.

It is difficult to say whether the three latent budgets that we found can be considered generic types of suicide. Heudin (1982,

TABLE 2b
Latent Budget Analysis Parameter Estimates for Data in Table 1
(Column Parameters $\hat{\pi}_{jt}^{\overline{B}X}$)

Cause of Death[a]	$\hat{\pi}_{+j}$	$t = 1$	$t = 2$	$t = 3$
1	.330	.530	.437	.000
2	.005	.011	.004	.000
3	.040	.135	.002	.000
4	.383	.000	.315	.840
5	.050	.021	.099	.015
6	.059	.104	.000	.090
7	.018	.007	.017	.029
8	.053	.052	.084	.016
9	.062	.141	.044	.010
Total	1.000	1.000	1.000	1.000

[a]See note to Table 1.

Chapter 7) points out that there are many possible explanations for the suicide method chosen. For example, it is argued that women care what they will look like after their death. This could explain their preference for the methods overrepresented in the second budget, which do not lead to mutilation of the body. In the third budget we find overrepresentation of methods that do mutilate the body. Another theory emphasizes the importance of opportunity. Some of the methods overrepresented in the first budget are methods for which one must have opportunity; i.e., one needs medicine, a car, or a gun (which is not easy to obtain in Germany). Younger people may have more opportunity to use these methods. Psychoanalytic theories suggest that suicides represent sexual wish fulfillments: To poison oneself is to become pregnant; to drown is to bear a child; to throw oneself from a height is to be delivered of a child. Poisoning, drowning, and jumping are three methods that are overrepresented in the second latent budget. There are also individual factors involved: Psychoanalysts suggest that suicide is a message to relatives. All the above explanations are highly speculative, although evidence of regional differences in methods used supports the opportunity theory. For example, people in the United States are more likely to commit suicide by using guns than people in countries where guns are not easily obtained. More information is needed to understand the relation of age and sex to the method chosen.

2.3. *Comparison with Latent Class Analysis*

Clogg (1981) presented latent budget analysis as a reparameterization of latent class analysis, and as a result many properties of latent class analysis also hold for latent budget analysis. Let π_t^X be the probability of falling into latent class t; let $\pi_{it}^{\bar{A}X}$ be the conditional probability of falling into level i of variable A given latent class t; and let $\pi_{jt}^{\bar{B}X}$ be the conditional probability of falling into level j of variable B given latent class t. Then the latent class model for two-way contingency tables is defined as a model for the latent probabilities π_{ijt} of falling into level i of variable A, level j of variable B, and level t of the latent variable X:

$$\pi_{ijt} = \pi_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X}. \tag{3a}$$

And the latent probabilities are related to the probabilities using only the manifest variables A and B by

$$\pi_{ij} = \sum_{t=1}^{T} \pi_{ijt}. \tag{3b}$$

Latent budget analysis and latent class analysis have in common the parameters $\pi_{jt}^{\bar{B}X}$. The latent budget analysis parameters $\pi_{it}^{A\bar{X}}$ are derived from the latent class analysis parameters π_t^X and $\pi_{it}^{\bar{A}X}$ by using Bayes's rule:

$$\pi_{it}^{A\bar{X}} = \frac{\pi_t^X \pi_{it}^{\bar{A}X}}{\sum_{t=1}^{T} \pi_t^X \pi_{it}^{\bar{A}X}}. \tag{4}$$

Latent budget analysis and latent class analysis are also readily compared in terms of the latent probabilities π_{ijt}. For both models the column parameters are related to the latent probabilities π_{ijt} as $\pi_{jt}^{\bar{B}X} = \pi_{+jt}/\pi_{++t}$. For latent class analysis the row parameters are related to the latent probabilities π_{ijt} as $\pi_{it}^{\bar{A}X} = \pi_{i+t}/\pi_{++t}$, whereas for latent budget analysis the row parameters are related to the latent probabilities π_{ijt} as $\pi_{it}^{A\bar{X}} = \pi_{i+t}/\pi_{i++}$. In terms of the latent probabilities, the observed variables A and B are conditionally independent given the latent variable X for both models. From (3a) it follows that latent class analysis can be defined as

$$\pi_{ijt} = \pi_{++t}\left(\frac{\pi_{i+t}}{\pi_{++t}}\right)\left(\frac{\pi_{+jt}}{\pi_{++t}}\right) = \frac{\pi_{i+t}\pi_{+jt}}{\pi_{++t}}. \tag{5}$$

Latent budget analysis implies that

$$\frac{\pi_{ijt}}{\pi_{i++}} = \left(\frac{\pi_{i+t}}{\pi_{i++}}\right)\left(\frac{\pi_{+jt}}{\pi_{++t}}\right) = \left(\frac{1}{\pi_{i++}}\right)\left(\frac{\pi_{i+t}\pi_{+jt}}{\pi_{++t}}\right).$$

Both the latent class model and the latent budget model can therefore be understood as a loglinear model with latent variables (see, e.g., Haberman 1979; Hagenaars 1986, 1988, 1990). This relation with loglinear analysis will be taken up later in the paper.

Latent class analysis is most often used to analyze contingency tables with more than two variables, where it aims to identify a latent variable that can explain the relations between the observed variables (see, e.g., Goodman 1974). Relatively little attention has been given to latent class analyses of two-way contingency tables, except in the theoretical contributions of Good (1969), Gilula (1979, 1983, 1984), Clogg (1981), Goodman (1987), and de Leeuw and van der Heijden (1991), and in the social mobility research of Marsden (1985), Grover (1987), Grover and Srinivasan (1987), and Luijkx (1987). As far as we know, the reparameterization (1) appears only in Clogg (1981) and in the references given in the introduction.

A choice between the latent class model and the latent budget model will depend on the research question at hand. If the question is about dependence, that is, how response variables depend on explanatory variables, then the latent budget model is more appropriate (see also section 2.4). If the question is about relations between variables, then latent class analysis is preferable. In this sense the distinction between latent class analysis and latent budget analysis is similar to the distinction between loglinear analysis and (multinomial) logit analysis. It will become apparent in section 3.2 that the distinction between explanatory variables and response variables suggests specific types of models for latent budget analysis that have not been considered thus far in the context of latent class analysis.

2.4. *Graphical Representations of the Model*

Figures 1 and 2 illustrate the usefulness of latent budget analysis with representations inspired by those used in covariance struc-

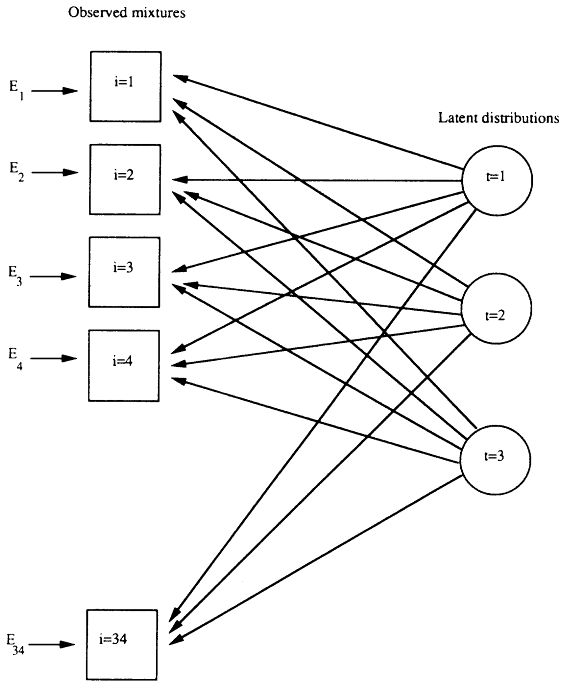

FIGURE 1. Latent budget model represented as a mixture model. The observed budgets are derived from three latent budgets plus error (E). Lines in the figure symbolize the mixing parameters

ture analysis. We have chosen two such representations, one emphasizing the interpretation as a mixture model, and one emphasizing the interpretation as a MIMIC model. In both, squares indicate observed entities, and circles latent entities.

In Figure 1, the latent budget model for Table 1 is represented as a mixture model. The observed contingency table is conceived as a matrix of I observed distributions (budgets), one for each row. These I observed distributions are generated by T latent distributions. The observed distributions have elements p_{ij}/p_{i+}, and the latent distributions have elements $\hat{\pi}_{jt}^{\overline{B}X}$. The 34 observed age-sex distributions of suicide types are represented by squares, and the latent distributions are represented by circles. The direction of the arrows in Figure 1 suggests that the mixtures of three latent distributions generate the observed distributions. The error E_i associated with each observed distribution represents the difference between the observed distribu-

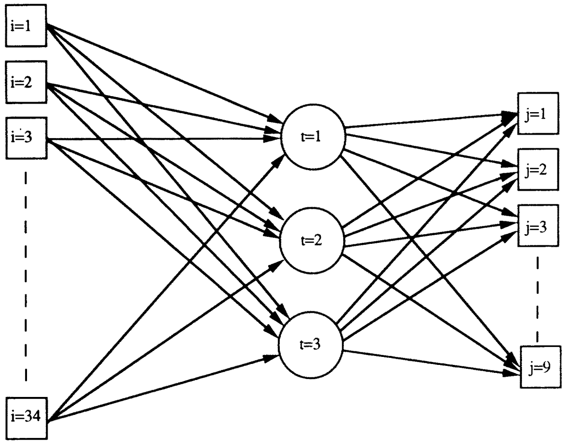

FIGURE 2. MIMIC model representation of the latent budget model.

tion elements p_{ij}/p_{i+} and the estimates of expected distribution elements $\hat{\pi}_{ij}/\hat{\pi}_{i+}$. Next to the arrows we could plot the parameter estimates $\hat{\pi}_{it}^{A\bar{X}}$, since they show how the mixture of the three latent budgets generates the estimates of the expected distributions. In the example in section 2.2, the latent distributions can be understood as typical distributions of cause-of-death categories, and our analysis has demonstrated that the observed distributions of the 34 age groups are relatively well approximated by three such typical distributions.

In Figure 2, the latent budget model is represented as a MIMIC model. In this case, the squares and circles do not denote distributions but categories of variables. This interpretation is due to Clogg (1981), who used it to represent a social mobility example. Here, it demonstrates that given a specific age-sex combination i ($i = 1, \ldots, 34$), there are probabilities associated with each of the three latent states t ($t = 1,2,3$); in each latent state there are nine probabilities ($j = 1, \ldots, 9$) corresponding to the nine cause-of-death categories. Now the estimates for the row parameters $\hat{\pi}_{it}^{A\bar{X}}$ as well as the estimates for the column parameters $\hat{\pi}_{jt}^{\bar{B}X}$ can be specified next to the arrows. The parameter estimates $\hat{\pi}_{it}^{A\bar{X}}$ can be placed beside the arrows from the age-sex group categories to the latent states. Beginning from a specific age-sex group, these estimates add up to one, showing that the probabilities associated with the latent states add up to one. The parameter estimates $\hat{\pi}_{jt}^{\bar{B}X}$ can be placed beside the arrows from the latent states to the cause-of-death categories. Starting from a given latent state, these estimates also add up to one, showing that the probability of falling in one of the cause-of-death categories is one. In the latent budget model in Figure 2, the squares and circles represent categories. In the usual covariance structure model, they represent *variables*. It is sometimes difficult to interpret the latent states. In the suicide example a latent state might stand for the experiences and influences that people have undergone, which might differ by age-sex group.

A latent class model would be similar to the model in Figure 2, except that the arrows would go from the latent classes to the row categories, and not from the row categories to the latent classes, as in the latent budget model. Latent budget analysis is the equivalent of latent class analysis, but the former is used to study asymmetric relations between variables and the latter is used to study symmetric relations. In this sense, the equivalence between latent class and latent budget analysis is similar to the equivalence between loglinear analysis and the (multinomial) logit model, the former intended in the first place for the study of symmetric relations between variables, the second for the study of asymmetric relations between variables.

2.5. *Identification*

For two observed variables, neither the unconstrained latent budget model nor the latent class model is identified, although the identification problem is rather well understood. For a detailed discussion of this identification problem in latent class analysis we refer to Goodman (1987); for latent budget analysis the problem is discussed in detail in de Leeuw et al. (1990). The latent budget analysis parameter estimates can be varied within a specific range without changing the estimates of expected frequencies for the model, suggesting enormous freedom for the parameter estimates. However, this freedom is smaller than it seems. The reason is that all parameter estimates covary, which means that interpretation remains rather stable over different choices of identifying restrictions for the model.

The identification problem is most easily explained by formalizing the latent budget model in matrix terms. Let $\mathbf{D}_r$ be a diagonal $I \times I$ matrix with marginal probabilities π_{i+} as elements; and let $\boldsymbol{\Pi}$ be the $I \times J$ matrix with probabilities π_{ij}. Collect the parameters $\pi_{it}^{A\bar{X}}$ in a $I \times T$ matrix **A**, and collect the parameters $\pi_{jt}^{\bar{B}X}$ in a $J \times T$ matrix **B**. Let **S** be a $T \times T$ square matrix with the property that $\mathbf{S1} = \mathbf{1}$, where **1** is a unit column vector of length T. Let $\mathbf{A}^* = \mathbf{AS}$ and $\mathbf{B}^* = \mathbf{B}(\mathbf{S}^{-1})'$ be matrices with alternative parameters. Then the identification problem can be written as

$$\mathbf{D}_r^{-1}\boldsymbol{\Pi} = \mathbf{AB}' = (\mathbf{AS})(\mathbf{S}^{-1}\mathbf{B}') = \mathbf{A}^*\mathbf{B}^{*\prime}. \qquad (6)$$

Since $\mathbf{S1} = \mathbf{1}$, the elements of $\mathbf{A}^*$ add up to 1 for each row and the elements of $\mathbf{B}^*$ add up to 1 for each column. **S** has to be chosen in such a way that all elements of $\mathbf{A}^*$ and $\mathbf{B}^*$ are non-negative, since these elements are probabilities.

The identification problem is very similar to the rotation problem in factor analysis, except that the abundance of solutions known for factor analysis is not yet available for latent budget analysis. De Leeuw et al. (1990) discuss some choices of **S** and suggest a way to obtain them. One possibility is to choose a matrix **S** that results in as many zeros in $\mathbf{A}^*$ or $\mathbf{B}^*$ as possible. This number of zeros is maximally $T(T - 1)$, the number of free elements of **S**. If we choose as many zeros in $\mathbf{A}^*$ as possible, interpretation is simplified, because if a parameter estimate $\hat{\pi}_{it}^{A\bar{X}}$ equals zero, then the estimates of the theoretical budget for row i are derived from less than T latent budgets. Of course, such a restriction would have to be sociologically justified. Many other details on the identification problem of the latent budget model can be found in de Leeuw et al. (1990).[1]

In the analysis of the suicide data the identification problem is negligible: Many row and column parameters are estimated at zero; therefore, no admissible matrix **S** will have a noticeable effect upon the parameter estimates (see de Leeuw et al. 1990).

The problem of identification is related to the derivation of the number of degrees of freedom:

$$df = \text{No. of nonredundant cells} - \text{No. of independent parameters.}$$

[1]This approach to identification emphasizes a choice of **S**, because by fitting the model with the EM algorithm, unidentified estimates **A** and **B** are obtained and are then transformed to some simple structure. It is also possible to start the other way around, i.e., to fix some parameters in **A** and **B**, try to prove that **S** can only be the identity matrix, and if this is the case, estimate the free parameters. The procedure advocated here has the advantage that the unidentified estimates suggest which elements of **A** and **B** can be fixed to zero without decreasing the fit of the model.

Given that the row totals n_{i+} are fixed, the number of nonredundant cells is $I(J-1)$. In the unconstrained case with T latent budgets, the number of row parameters is $I(T-1)$, and the number of column parameters is $(J-1)T$. However, because these parameters are not independent, since $\mathbf{AB}' = \mathbf{ASS}^{-1}\mathbf{B}'$, we must subtract the total number of free elements of $\mathbf{S}$ from the estimated parameters. This number is $T(T-1)$. Hence, in the unconstrained case we get

$$df = I(J-1) - [I(T-1) + (J-1)T - T(T-1)] = (I-T)(J-T).$$

2.6. *Maximum Likelihood Estimation*

Here we discuss the estimation of the model. Readers who are not interested in this subject can skip this section without loss of continuity.

We estimate the model with the EM algorithm (Dempster, Laird, and Rubin 1977), which is also the algorithm most often employed for the estimation of latent class models (see Goodman [1974] for a description of the algorithm). Alternative algorithms used for the estimation of latent class models are provided by Haberman (1979, 1988) and by Formann (1978).

The EM algorithm is used to estimate missing values. In this paper the observations on the latent variable X are missing, and only the marginal frequencies n_{ij} of the three-way matrix with elements n_{ijt} are observed. For the unobserved three-way matrix with probabilities π_{ijt}, the latent budget model (1) is defined in terms of π_{ijt}/π_{i++} by $\pi_{ijt}/\pi_{i++} = \pi_{it}^{A\overline{X}}\pi_{jt}^{\overline{B}X}$. The loglikelihood for the unobserved matrix is

$$L = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T} n_{ijt} \ln \frac{\pi_{ijt}}{\pi_{i++}}. \tag{7}$$

The unobserved n_{ijt} and the parameter estimates for $\pi_{it}^{A\overline{X}}$ and $\pi_{jt}^{\overline{B}X}$ are unknown.

The EM algorithm consists of two steps: the expectation step (E-step) and the maximization step (M-step). In the E-step the expectation of the loglikelihood for the unobserved data n_{ijk} is found, conditional on the observed frequencies n_{ij} and the current parameter estimates. The expectations of the sufficient statistics of the complete data matrix therefore have to be expressed in terms of the model parameters, so we need an expression for n_{ijt}. For this step the current best estimates of $\pi_{it}^{A\overline{X}}$ and $\pi_{jt}^{\overline{B}X}$ are taken. Thus, updated estimates of the unobserved frequencies $\underline{n}_{ijt}$ are given by

$$\underline{n}_{ijt} = n_{ij}\frac{\pi_{ijt}}{\pi_{ij+}} = n_{ij}\frac{\pi_{ijt}/\pi_{i++}}{\pi_{ij}/\pi_{i+}} = n_{ij}\left(\frac{\pi_{it}^{A\overline{X}}\pi_{jt}^{\overline{B}X}}{\sum_{t=1}^{T}\pi_{it}^{A\overline{X}}\pi_{jt}^{\overline{B}X}}\right). \tag{8}$$

In the M-step the loglikelihood for the unobserved data is maximized as a function of the model parameters. Let γ_i and δ_t be Lagrange multipliers needed for the constraints (2). Then the following function is to be maximized over $\pi_{it}^{A\overline{X}}$ and $\pi_{jt}^{\overline{B}X}$:

$$\begin{aligned} f(\pi_{it}^{A\overline{X}},\pi_{jt}^{\overline{B}X},\gamma_i,\delta_t) = & \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T}\underline{n}_{ijkt}\ln(\pi_{it}^{A\overline{X}}\pi_{jt}^{\overline{B}X}) \\ & -\sum_{i=1}^{I}\gamma_i\Big((\sum_{t=1}^{T}\pi_{it}^{A\overline{X}})-1\Big)-\sum_{t=1}^{T}\delta_t\Big((\sum_{j=1}^{J}\pi_{jt}^{\overline{B}X})-1\Big). \end{aligned} \tag{9}$$

To find the updated estimates $\underline{\pi}_{it}^{A\overline{X}}$ and $\underline{\pi}_{jt}^{\overline{B}X}$ that maximize $f(\pi_{it}^{A\overline{X}},\pi_{jt}^{\overline{B}X},\gamma_i,\delta_t)$, we can rewrite (9) as both (10a) and (10b) and maximize these over $\pi_{it}^{A\overline{X}}$ and $\pi_{jt}^{\overline{B}X}$ separately:

$$f(\pi_{it}^{A\overline{X}},\gamma_i) = \sum_{i=1}^{I}\sum_{t=1}^{T}\underline{n}_{i+t}\ln\pi_{it}^{A\overline{X}} - \sum_{i=1}^{I}\gamma_i\left((\sum_{t=1}^{T}\pi_{it}^{A\overline{X}})-1\right) + \text{constant}, \tag{10a}$$

$$f(\pi_{jt}^{\overline{B}X},\delta_t) = \sum_{j=1}^{J}\sum_{t=1}^{T}\underline{n}_{+jt}\ln\pi_{jt}^{\overline{B}X} - \sum_{t=1}^{T}\delta_t\left((\sum_{j=1}^{J}\pi_{jt}^{\overline{B}X})-1\right) + \text{constant}. \tag{10b}$$

This gives as updated estimates $\underline{\pi}_{it}^{A\overline{X}} = \underline{n}_{i+t}/\underline{n}_{i++}$ and $\underline{\pi}_{jt}^{\overline{B}X} = \underline{n}_{+jt}/\underline{n}_{++t}$. These updated estimates are used in (8) as current best estimates in the next E-step of the algorithm.

Initial parameter estimates should be consistent with the constraints (2). If initial estimates equal to zero are chosen, their values will not change throughout the algorithm. De Leeuw et al. (1990) prove that in this application of the EM algorithm, the likelihood increases in each step. Therefore, the algorithm converges. To investigate whether it converges to a global maximum, one should try different sets of initial estimates.

3. CONSTRAINING THE PARAMETERS

In this section we will describe and illustrate three types of constraints on parameters: fixed-value constraints, equality constraints, and multinomial logit constraints. Since latent budget analysis is closely related to latent class analysis, the procedure for constraining the parameters in latent budget analysis is similar to the procedure in latent class analysis. Fixed-value constraints and equality constraints are well known in that context, so they receive less attention in this paper. Langeheine (1989) gives an insightful overview of constraints in latent class analysis and relates the types of constraints to the different ways latent class analysis is presented—for example, as a product of conditional probabilities, as in (3a) and (3b) (Goodman 1974), and as a loglinear model with a latent variable (Haberman 1979). We use both representations in our discussion of constraints in latent budget analysis.

3.1. *Fixed-Value and Equality Constraints*

Theory. A fixed-value constraint on a parameter fixes this parameter to a prespecified value, for example, $\pi_{it}^{A\overline{X}} = c$, or $\pi_{jt}^{\overline{B}X} = c'$, where $0 \leq c, c' \leq 1$. It is important here to distinguish between identifying constraints and constraints that affect the fit of the model. In section 2.5, we chose a matrix **S** to resolve the identification problem. This choice of a specific **S** implies a choice of a specific solution, for example, a solution with as many parameters $\pi_{it}^{A\overline{X}}$ as possible fixed to zero, or as many parameters $\pi_{jt}^{\overline{B}X}$ as possible fixed to zero (see de Leeuw et al. 1990). Such fixed parameters are called identifying constraints. To impose these constraints, we first identify the model by choosing a matrix **S**, thus implicitly constraining some of the parameters to certain values. We can then constrain one or more further parameters. The model with both the identifying constraints and the extra constraints can be estimated by considering both types of constrained parameters as fixed parameters.

Fixed-value constraints are useful for testing whether parameter estimates differ significantly from values that are of theoretical interest. In many circumstances such values will be zero or one. For the suicide data discussed in section 2, additional constraints could be imposed upon the parameters $\pi_{jt}^{\overline{B}X}$ for budget $t = 3$, for example, by constraining all probabilities to zero except those for hanging, stabbing, and shooting deaths. Thus, the third budget would be a budget consisting only of more-violent causes of death. Subsequently, the row parameters $\pi_{it}^{A\bar{X}}$ for women in the third latent budget could be constrained to zero, so that only males were in the third latent budget. Such fixed-value constraints would simplify the interpretation considerably.

An interesting example of a fixed-value constraint is $\pi_{j1}^{\bar{B}X} = \ldots = \pi_{jt}^{\bar{B}X} = \ldots = \pi_{jT}^{\bar{B}X} = p_{+j}$; that is, element j of each latent budget is equal to the sample proportion that falls into category j. Thus, for all i and some category j, $\pi_{ij}/\pi_{i+} = \pi_{+j}$. This allows us to determine whether the differences between the budgets π_{ij}/π_{i+} are due to differences in column categories other than column category j. This hypothesis can be tested against the unconstrained model with T latent classes; since T parameters are constrained, the test statistic is asymptotically chi-square distributed with T degrees of freedom. If such a hypothesis cannot be rejected, the interpretation is simplified considerably: It is no longer necessary to characterize the groups (rows) in terms of differences in their use of column j. Equality constraints specify that certain row-parameter estimates $\pi_{it}^{A\bar{X}}$ or certain column-parameter estimates $\pi_{jt}^{\bar{B}X}$ are unknown but equal to one another. Such equality constraints can be used to test whether two parameter estimates really are different. If this is the case, the interpretation is again simplified considerably. As with fixed-value constraints, the model should first be identified by some choice of **S**. Then certain parameter estimates can be constrained to be equal.

An interesting test for equality constraints amounts to a test for the collapsibility of rows. If $\pi_{it}^{A\bar{X}} = \pi_{i't}^{A\bar{X}}$ for all t, then rows i and i' are related to the latent budgets in the same way. This implies for the theoretical budget elements in rows i and i' that $\pi_{ij}/\pi_{i+} = \pi_{i'j}/\pi_{i'+}$. The equality of theoretical row budgets was used earlier by Goodman (1981), Gilula (1986), and Gilula and Krieger (1989) as a criterion for the collapsibility of rows. In a similar way we can also interpret the test that $\pi_{it}^{A\bar{X}} = \pi_{i't}^{A\bar{X}}$ for all t as a test for collapsibility of rows i and i'. Gilula (1986) and Gilula and Krieger (1989) applied similar tests for equality of the parameters of two or more rows (or two or more columns) in a maximum likelihood version of correspondence analysis. The similarity between the work of Gilula and the procedure described here follows immediately in all cases in which correspondence analysis and latent budget analysis are equivalent (see de Leeuw and van der Heijden 1991). In cases in which correspondence analysis and latent budget analysis are not equivalent, the tests can give different results.

Derivation of the number of degrees of freedom for the unconstrained case was discussed above in section 2.5. In the unconstrained case we usually choose an **S** that gives us as many zero parameters in **A** or **B** as possible, thus simplifying the interpretation. This can also be done by constraining these estimates to zero. Each additional fixed-value constraint leads to a higher number of degrees of freedom. Similar remarks apply to equality constraints. First, we constrain $T(T-1)$ parameters so that the model is identified. Then, after imposing equality constraints, we can easily derive the number of degrees of freedom.

We will not discuss examples of fixed-value or equality constraints. Some applications are suggested above, and examples can be found in the latent class analysis literature (see, e.g., McCutcheon 1987).

Estimation.[2] For the constraints discussed above, the loglikelihood can always be split into two parts, as in (10a) and (10b), which can be maximized separately to update parameter estimates.

We first introduce some notation. We denote parameters constrained to fixed values by adding a bar over the parameter symbol, that is, $\bar{\pi}_{it}^{A\bar{X}}$ and $\bar{\pi}_{jt}^{\bar{B}X}$. In the presence of fixed-value constraints, we denote free parameters by adding a tilde over the parameter symbol, that is, parameters to be estimated are $\tilde{\pi}_{it}^{A\bar{X}}$ and $\tilde{\pi}_{jt}^{\bar{B}X}$. From the parameters $\bar{\pi}_{it}^{A\bar{X}}$, $\bar{\pi}_{jt}^{\bar{B}X}$, $\tilde{\pi}_{it}^{A\bar{X}}$, and $\tilde{\pi}_{jt}^{\bar{B}X}$ we derive the elements $\pi_{it}^{A\bar{X}}$ and $\pi_{jt}^{\bar{B}X}$ of matrices **A** and **B**. For this purpose the current estimates of the parameters $\tilde{\pi}_{it}^{A\bar{X}}$ and $\tilde{\pi}_{jt}^{\bar{B}X}$ are sometimes rescaled so that estimates for the parameters $\pi_{it}^{A\bar{X}}$ and $\pi_{jt}^{\bar{B}X}$ follow constraints (2).

Fixed-value constraints for specific row and column parameters can be easily imposed. This is shown for fixed row parameters only. (The results for the column parameters can be derived in a similar way.) Let the parameter for row i in budget m be fixed to some constant c; that is, $\bar{\pi}_{im}^{A\bar{X}} = c, 0 \leq c \leq 1$. Since the elements $\pi_{it}^{A\bar{X}}$ of matrix **A** are constrained by (2), we have

$$\pi_{it}^{A\bar{X}} = (1 - \sum_m \bar{\pi}_{im}^{A\bar{X}})\tilde{\pi}_{it}^{A\bar{X}}, \quad \text{with } \sum_j \tilde{\pi}_{it}^{A\bar{X}} = 1, \tag{11}$$

[2]This section can be skipped without loss of continuity.

where $\Sigma_m \bar{\pi}_{im}^{A\bar{X}}$ implies that the sum is taken over the fixed parameters for row i, and $\Sigma_t \tilde{\pi}_{it}^{A\bar{X}}$ implies that the sum is taken over the free parameters for row i. To estimate the elements $\pi_{it}^{A\bar{X}}$ of matrix **A**, we maximize the loglikelihood in the M-step over the parameters $\tilde{\pi}_{it}^{A\bar{X}}$:

$$f(\tilde{\pi}_{it}^{A\bar{X}}, \gamma_i) = \sum_{\substack{\text{index pairs } (i,t) \\ \text{corresponding with} \\ \text{free parameters}}} \left[n_{i+t} \ln \left(\left(1 - \sum_m \bar{\pi}_{im}^{A\bar{X}} \right) \tilde{\pi}_{it}^{A\bar{X}} \right) \right] - \sum_{i=1}^{I} \gamma_i \left(\left(\sum_j \tilde{\pi}_{it}^{A\bar{X}} \right) - 1 \right) + \text{constant}. \tag{12}$$

Maximizing (12) over $\tilde{\pi}_{it}^{A\bar{X}}$, we find the estimate $\tilde{\pi}_{it}^{A\bar{X}} = n_{i+t}/n_{i++}$, where in obtaining n_{i++} and n_{i+t}, we use only those frequencies with index pairs (i,t) of free parameters. Having $\tilde{\pi}_{it}^{A\bar{X}}$ we can now find new estimates $\pi_{it}^{A\bar{X}}$ using (11).

We next consider both equality constraints and fixed-value constraints. A bar over the parameter symbol indicates that the parameter is again fixed to some specific value. A tilde indicates that the parameter is either completely free or constrained to be equal to other unknown parameters. These two types of parameter constitute the matrices **A** with elements $\pi_{it}^{A\bar{X}}$ and **B** with elements $\pi_{jt}^{\bar{B}X}$. We consider only equality constraints between row parameters and equality constraints between column parameters. We do not consider equality constraints between row and column parameters. Thus, the estimation problem in the M-step can be simplified because we can rewrite the loglikelihood, separating the parameters for **A** from the parameters for **B** and maximizing these parameters separately (compare (9), (10a), and (10b)).

We first consider constraints on row parameters $\pi_{it}^{A\bar{X}}$. Equation (12) represents the loglikelihood function to be maximized over the free parameters $\tilde{\pi}_{it}^{A\bar{X}}$, where there are only fixed-value constraints and no equality constraints. The updated estimates $\tilde{\pi}_{it}^{A\bar{X}}$ can be transformed into their corresponding $\pi_{it}^{A\bar{X}}$ via (11). Now let some parameters $\pi_{it}^{A\bar{X}}$ be restricted to be equal to one another. We introduce some extra notation. Let A_{it} be the set of pairs $\{(i,t)\}$ for which equality constraints are imposed; we denote this set by the first combination (i,t) encountered (the index i running faster than the index t). That is, if $\pi_{12}^{A\bar{X}} = \pi_{21}^{A\bar{X}}$, then the set $\{(1,2),(2,1)\}$ is A_{12}, and A_{21} is not defined.

If no equality constraint is imposed for some π_{it}^{AX}, then A_{it} has as its only element (i,t). Let W be the set of existing index pairs (i,t) of the sets A_{it}. Let the number of parameters in each A_{it} be equal to c_{it}, and let $f_{it} = \Sigma_{\text{index pairs in } A_{it}} n_{i+t}$. Instead of (12) we find that

$$f(\tilde{\pi}_{it}^{AX}, \gamma_i) = \sum_{\substack{\text{index pairs } (i,t)\\ \text{in } W}} \left[f_{it} \ln \left((1 - \sum_m \tilde{\pi}_{im}^{AX}) \tilde{\pi}_{it}^{AX} \right) \right] - \sum_{i=1}^{I} \gamma_i \left((\sum_t c_{it} \tilde{\pi}_{it}^{AX}) - 1 \right) + \text{constant}. \tag{13}$$

Now we maximize (13) over $\tilde{\pi}_{it}^{AX}$ and use the updated estimate $\tilde{\pi}_{it}^{AX}$ in (11) to find the updated parameter estimates π_{it}^{AX}. For example, if the equality constraint is $\pi_{11}^{AX} = \pi_{12}^{AX}$, then the updated estimate found for π_{11}^{AX} is $\pi_{11}^{AX} = (n_{1+1} + n_{2+1})/(n_{1++} + n_{2++})$, and the other parameters in rows 1 and 2 may be derived from n_{i+t}/n_{i++}, then adjusted in a way similar to (11) so that (2) holds. In most cases the maximization of (13) over $\tilde{\pi}_{it}^{AX}$ gives direct solutions. There are cases, however, in which no direct solutions exist. In such a case, a Newton-Raphson or similar procedure should be applied in each M-step of the EM algorithm. This can make the algorithm very time-consuming. For a discussion of the existence of direct solutions in latent class analysis, see Mooijaart and van der Heijden (in press). Since latent class analysis and latent budget analysis are equivalent, their results can readily be applied to the latent budget model.

3.2. *Multinomial Logit Constraints*

If there is additional information about the rows and columns of the two-way table, it is possible to constrain the corresponding row and column parameters as a function of this additional information. One sort of additional information specifies that the row variable or the column variable is a joint variable. The suicide data analyzed in section 2 have a joint row variable, based on the separate variables age and sex. In section 3.5 we show how this type of additional information about the rows can be used in the model. Another sort of information that we can use is any available quantitative information relevant to the row or column categories. For example, consider a matrix of industry groups by years, in which the number of firms in each industry founded in each year are in the cells. The observed budgets give the proportions of types of firms founded in each year, and the latent budgets give typical distributions of new firms by industry. If the row categories of a matrix are specific years, then these years themselves could be used as additional information in the model. It is also possible to use other variables that describe the economic situation in these years to model the differences between the observed budgets.

The parameters $\pi_{it}^{A\bar{X}}$ and $\pi_{jt}^{\bar{B}X}$ are conditional probabilities; therefore, we use the multinomial logit model, which has been devised specifically for this situation. We use the version of the multinomial logit model discussed extensively by Bock (1975) to constrain both the row and column parameters.

The multinomial logit model has been used to constrain parameters in other contexts. Formann (1982, 1985, 1989) constrains latent class analysis in a manner very similar to our own but concentrates on latent class analysis of dichotomous variables (see also Langeheine 1989). Shigemasu and Sugiyama (1989) use this parameterization for latent class analysis of choice behavior. Takane (1987) uses the multinomial logit model to restrict the conditional probabilities in ideal point discriminant analysis.

Multinomial logit constraints on row parameters. Let the additional information for the rows i of the two-way contingency table be collected in an $I \times M$ matrix $\mathbf{V}$, where the columns are indexed by m ($m = 1, \ldots, M$). We can use this information by defining the following model for the (conditional) row parameters $\pi_{it}^{A\bar{X}}$.

$$\pi_{it}^{A\bar{X}} = \frac{\exp\left(\sum_{m=1}^{M} v_{im}\gamma_{mt}\right)}{\sum_{t=1}^{T}\exp\left(\sum_{m=1}^{M} v_{im}\gamma_{mt}\right)}. \tag{14}$$

Here, γ_{mt} is the (m,t)th element of the $M \times T$ parameter matrix $\boldsymbol{\Gamma}$. Model (14) can be identified by constraining $\gamma_{m1} = 0$, for all m.

Model (14) allows a large range of constraints to be defined by the matrix $\mathbf{V}$. We get more insight into this class of constraints by relating unconstrained latent budget analysis to loglinear analysis with latent variables. Let π_{ijt} be the latent probabilities. Then equation (5) shows that variables A and B are conditionally independent given the level of the latent variable X. Thus, using notation similar to that employed by Haberman (1979, Chapter 10), we get

$$\log \pi_{ijt} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_t^X + \lambda_{it}^{AX} + \lambda_{jt}^{BX}. \tag{15}$$

From (14), the conditional probabilities $\pi_{i+t}/\pi_{i++} = \pi_{it}^{A\bar{X}}$ are further constrained. Thus, although the variables A and B remain conditionally independent given the level of the latent variable X, the interaction between A and X, λ_{it}^{AX}, is further constrained.

An important special case of (14) is when the row variable is itself a cross-classification of two or more variables, and the matrix $\mathbf{V}$ is a design matrix describing the factorial structure of the rows. In this case, we can define a multinomial logit model for the row parameters $\pi_{it}^{A\bar{X}}$ that is equivalent to a hierarchical loglinear model for unconditional marginal probabilities $\pi_{i+}t$. The hierarchical loglinear model to which it is equivalent has the same sets of parameters, which describe the relations between the explanatory and the response variables, as the multinomial logit model and additional sets of parameters that describe the relations between the explanatory variables (see Bock 1975; Haberman 1979). This is well known in the case of two latent budgets ($T = 2$), since then the multinomial logit model simplifies to the ordinary logit model (see, for example, Fienberg 1980, Chapter 6). In fact, if the loglinear parameter estimates are identified in the same way that the multinomial logit parameter estimates were identified above (i.e., by setting the first parameter of each set equal to zero), then estimates for the multinomial logit parameters are equal to estimates for the corresponding loglinear parameters.

The equivalence of the multinomial response model and the hierarchical loglinear model gives us further insight into the constrained latent budget model. Goodman (1971) showed that if the marginal probabilities $\pi_{i+}t$ follow a restrictive hierarchical loglinear model including parameters for t (which correspond to a column of 1's in $\mathbf{V}$), and if for the latent probabilities π_{ijt} the variables A and B are independent given the level of variable X, then the probabilities π_{ijt} still follow a hierarchical loglinear model. It follows from collapsibility properties in loglinear models that have conditional independence properties (see Whittaker 1990). Thus, constraining the latent budget model does not affect the relationship between latent budget analysis and loglinear analysis: The constrained latent budget model can still be considered a hierarchical loglinear model for the latent probabilities.

Consider an example. Let the rows of a two-way contingency table be stratified by two variables, A and C, with categories of A indexed by i and categories of C indexed by k ($k = 1, \ldots, K$). The latent budget model can be written as

$$\frac{\pi_{ikj}}{\pi_{ik+}} = \sum_{t=1}^{T} \pi_{ikt}^{AC\bar{X}} \pi_{jt}^{\bar{B}X}. \tag{16}$$

Consider the case in which the matrix $\mathbf{V}$ in (14) does not constrain the row parameters; that is, it specifies a saturated multinomial logit model. Then the latent loglinear model is (compare (15))

$$\begin{aligned} \log \pi_{ikjt} = \lambda &+ \left(\lambda_i^A + \lambda_k^C + \lambda_{ik}^{AC}\right) + \lambda_j^B + \lambda_t^X \\ &+ \left(\lambda_{it}^{AX} + \lambda_{kt}^{CX} + \lambda_{ikt}^{ACX}\right) + \lambda_{jt}^{BX}. \end{aligned} \tag{17}$$

We can constrain $\pi_{ikt}^{AC\bar{X}}$ by omitting the columns of the matrix $\mathbf{V}$ corresponding to the two-factor interactions λ_{it}^{AX} and λ_{kt}^{CX} or the three-factor interaction λ_{ikt}^{ACX}. For example, we could constrain λ_{ikt}^{ACX} to equal zero for all i, k, and t by deleting the columns of $\mathbf{V}$ that describe the interaction between i and k: If a model with this constraint fits adequately, we can conclude that the row budgets cross-classified by variables A and C can be adequately approximated in terms of the latent budgets $\pi_{jt}^{\bar{B}X}$ by an effect of A and an effect of C, although in terms of the latent loglinear model, there is no interaction between these variables in their relation to the latent budgets. Thus, the latent budget model can still be understood as a loglinear model for the latent probabilities.

Although the multinomial logit parameters γ_{mt} (or the corresponding λ parameters) are the fundamental parameters of the model defined for the conditional probabilities $\pi_{it}^{A\bar{X}}$, these fundamental parameters are difficult to interpret. Usually we do not study the γ_{mt} parameters. Often we do not even calculate them. Instead, we study the conditional probabilities $\pi_{it}^{A\bar{X}}$ that they yield. A drawback of interpreting the conditional probabilities instead of the fundamental parameters is that the constraints imposed by the set of fundamental parameters are not always clearly revealed by the conditional probabilities $\pi_{it}^{A\bar{X}}$ used in latent budget analysis. The properties that hold for $\log \pi_{ijt}$ are affected by the way in which the conditional probabilities $\pi_{it}^{A\bar{X}}$ are defined in (14). However, in situations like model (17) with constraint $\lambda_{ikt}^{ACX} = 0$, it is possible to study marginal probability estimates $\hat{\pi}_{it}^{A\bar{X}} = \hat{\pi}_{i++}t/\hat{\pi}_{i+++}$ and $\hat{\pi}_{kt}^{C\bar{X}} = \hat{\pi}_{+k+}t/\hat{\pi}_{+k++}$ instead of estimates $\hat{\pi}_{ikt}^{AC\bar{X}} = \hat{\pi}_{ik+}t/\hat{\pi}_{ik++}$ because the margins $\hat{\pi}_{i++}t$ and $\hat{\pi}_{+k+}t$ are two of the margins that generated the hierarchical loglinear model for the latent probabilities π_{ikjt}. Interpreting parameter estimates $\hat{\pi}_{it}^{A\overline{X}}$ and $\hat{\pi}_{kt}^{C\overline{X}}$ instead of $\hat{\pi}_{ikt}^{AC\overline{X}}$ can simplify the interpretation considerably.

To illustrate the usefulness of multinomial logit constraints for the row parameters, we will discuss two examples below.

Multinomial logit constraints on column parameters. Using the multinomial logit model, we can define constraints on the column parameters (see Bock 1975). We assume that the information on the column categories of the two-way contingency table is collected in the matrix **W**, which has J rows and H columns and is indexed by h ($h = 1, \ldots, H$). Let $\boldsymbol{\Psi}$ be an $H \times T$ matrix with parameters ψ_{ht}. Now we can constrain the $\pi_{jt}^{\overline{B}X}$ parameters by the multinomial logit model

$$\pi_{jt}^{\overline{B}X} = \frac{\exp\left(\sum_{h=1}^{H} w_{jh}\psi_{ht}\right)}{\sum_{n=1}^{J} \exp\left(\sum_{h=1}^{H} w_{nh}\psi_{ht}\right)} . \tag{18}$$

The matrix **W** defines a model that is fitted to each latent budget t separately.

If **W** is chosen in an appropriate way, then the multinomial logit model for $\pi_{jt}^{\overline{B}X}$ is equivalent to a hierarchical loglinear model for the marginal probabilities π_{+jt}. Goodman (1971) showed that if π_{+jt} follows a hierarchical loglinear model and variables A and B are independent given X, the latent probabilities π_{ijt} also follow a hierarchical loglinear model. The situation is therefore similar to that for multinomial logit constraints on the rows.

When the column categories are classified by more than one variable, we can use the matrix **W** to describe the factorial structure of the columns. In this situation, an important special case is a constrained version of the simultaneous latent class model. Let there be two variables for the column categories, B indexed by j and D indexed by m. Then the latent budget model is

$$\frac{\pi_{ijm}}{\pi_{i++}} = \sum_{t=1}^{T} \pi_{it}^{A\overline{X}} \pi_{jmt}^{\overline{BD}X} , \tag{19}$$

and the loglinear model for the latent probabilities π_{ijmt} is

$$\begin{aligned} \log\pi_{ijmt} = {} & \lambda + \lambda_i^A + \left(\lambda_j^B + \lambda_m^D + \lambda_{jm}^{BD}\right) \\ & + \lambda_t^X + \lambda_{it}^{AX} + \left(\lambda_{jt}^{BX} + \lambda_{mt}^{DX} + \lambda_{jmt}^{BDX}\right). \end{aligned} \tag{20}$$

Let **W** constrain the variables B and D to be independent at each level of X. This is achieved when **W** consists of columns contrasting the categories of variable B and the categories of variable D but there are no contrasts linking them. Then in the loglinear model for the marginal probabilities π_{+jmt}, B and D are conditionally independent given X (see Bock 1975). This yields model (20) with additional constraints $\lambda_{jm}^{BD} = \lambda_{jmt}^{BDX} = 0$. Model (19) can then be rewritten as

$$\frac{\pi_{ijm}}{\pi_{i++}} = \sum_{t=1}^{T} \pi_{it}^{A\bar{X}} \pi_{jt}^{\bar{B}X} \pi_{mt}^{\bar{D}X} . \tag{21}$$

Thus we have obtained a constrained form of simultaneous latent class analysis. In this constrained version of the simultaneous latent class model, variables B and D are conditionally independent given X, and the relationship between variables B and D is identical for every level of the group variable A. This identity makes (21) a *constrained* version of the simultaneous latent class model (cf. Clogg and Goodman 1984). The levels of the group variable A may be related in different ways to the latent budgets, however.

To illustrate the usefulness of multinomial constraints for the column parameters, we analyze the suicide data that was discussed in section 2.

Degrees of freedom and identification. By imposing multinomial logit constraints on row and column parameters, we sometimes identify the model. If the model is identified by imposing constraints as in (14) and (18), then the matrix **S** in $\mathbf{AB}' = \mathbf{ASS}^{-1}\mathbf{B}'$ can only be the identity matrix (cf. Mooijaart 1982), and we should not subtract $T(T - 1)$ from the number of estimated parameters. This occurs in many constrained models.

If there is uncertainty about the number of degrees of freedom, this number can be derived by standard methods, for example, by determining the rank of the matrix of partial derivatives of the probabilities with respect to the parameters (cf. Goodman 1974).

Maximum likelihood estimation. When there are multinomial logit models for both the row and column parameters, then a loglikelihood function that is similar to (9) can be defined, and as in (10a) and (10b), it can be split into two parts, one part for the row parameters and one part for the column parameters. These parts can be maximized separately. They are

$$f(\gamma_{tm}) = \sum_{i=1}^{I} \sum_{T=1}^{T} \underline{n}_{i+t} \ln \left(\frac{\exp\left(\sum_{m=1}^{M} v_{im}\gamma_{tm}\right)}{\sum_{n=1}^{T} \exp\left(\sum_{m=1}^{M} v_{im}\gamma_{nm}\right)} \right) + \text{constant} \quad (22a)$$

and

$$f(\psi_{th}) = \sum_{j=1}^{J} \sum_{t=1}^{T} \underline{n}_{+jt} \ln \left(\frac{\exp\left(\sum_{h=1}^{H} w_{jh}\psi_{th}\right)}{\sum_{n=1}^{J} \exp\left(\sum_{h=1}^{H} w_{nh}\psi_{th}\right)} \right) + \text{constant}, \quad (22b)$$

where (22a) is to be maximized over γ_{tm} and (22b) is to be maximized over ψ_{th}.

For the row parameters $\pi_{it}^{A\overline{X}}$, this amounts to fitting the multinomial logit model to the updated estimate of the margins of the latent frequencies $\underline{n}_{i+}t$, using these as if they were ordinary observed frequencies. For the column parameters $\pi_{jt}^{\overline{B}X}$, this amounts to fitting the multinomial logit model to the updated estimate of the margins of the latent frequencies $\underline{n}_{+jk}$, using these as if they were ordinary observed frequencies.

Bock (1975) shows how to find the estimates for the general multinomial logit model using the Newton-Raphson algorithm. He also explains how to fit the multinomial logit model in case of structural zeros, which can be used in our context whenever there are additional fixed-value constraints for some of the $\pi_{it}^{A\overline{X}}$.

We deal now with an important special case in which the multinomial logit model is equivalent to a hierarchical loglinear model. This might occur for the rows when the matrix **V** describes the factorial structure of the rows of the contingency table, and it might occur for the columns when the matrix **W** describes the factorial structure of the columns of the contingency table. In such cases, the multinomial logit model can be fitted using iterative proportional fitting (see Bishop, Fienberg, and Holland 1975; Fienberg 1980, Chapters 3 and 4). Iterative proportional fitting is computationally more efficient than the Newton-Raphson procedure either when the number of parameters to be estimated becomes large or when direct estimates for the expected frequencies exist.

When iterative proportional fitting is used to fit a multinomial logit model for the row parameters $\pi_{it}^{A\overline{X}}$, for example, then this procedure yields constrained estimates $\hat{\pi}_{it}^{A\overline{X}}$ but no estimates for the multinomial logit parameters γ_{mt} (compare (14)). However, the constrained estimates $\hat{\pi}_{it}^{A\bar{X}}$ can be used to obtain the multinomial logit parameters in the following way. Let the estimates of conditional probabilities $\hat{\pi}_{it}^{A\bar{X}}$ follow some multinomial logit model. In this case the unconditional estimates $p_{i+}\hat{\pi}_{it}^{A\bar{X}}$ follow a corresponding loglinear model. If the loglinear parameters are identified in the same way as the multinomial logit parameters, that is, by constraining the first of a set of parameters to be zero, then the estimates for the multinomial logit parameters γ_{mt} are identical to the corresponding estimates for the loglinear λ parameters. Thus, the estimates for the loglinear λ parameters can be derived easily from $\log p_{i+}\hat{\pi}_{it}^{A\bar{X}}$, since $\log p_{1+}\hat{\pi}_{11}^{A\bar{X}} = \hat{\lambda}$, $\log p_{1+}\hat{\pi}_{12}^{A\bar{X}} = \hat{\lambda} + \hat{\lambda}_2^X = \log p_{1+}\hat{\pi}_{11}^{A\bar{X}} + \hat{\lambda}_2^X$, and so on.

Examples. For the suicide data analyzed in section 2, we chose the unconstrained model with $T = 3$. The fit of this model was $G^2 =$ 1,085.9, $df = 186$, although the fit measures were used as descriptive measures only.

Let the row variable age be A, indexed by i, and let the row variable sex be C, indexed by k. Then the unconstrained latent budget model is (16) or, as the equivalent loglinear model for the latent probabilities π_{ikjt}, (17). As discussed above, we will use the factorial structure of the rows to constrain the row parameters $\pi_{ikt}^{AC\bar{X}}$. The matrix **V** used to constrain these parameters has as its first column a unit vector, as its second column a dummy vector for sex, and then 16 dummy vectors for age. The unconstrained model would have 16 extra dummy vectors for the interaction between sex and age. We will investigate here whether this interaction is important by omitting these sixteen columns. In terms of the loglinear model for the latent probabilities π_{ikjt}, this corresponds to constraining the parameters $\lambda_{ikt}^{AC\bar{X}}$ to be zero.

This constraint can be motivated as follows. In the unconstrained model the expected budgets of the age-sex groups are related to three latent budgets. There can be a sex effect, an age effect, and an age-sex interaction effect. If the interaction effect can be omitted, then the interpretation simplifies considerably: Only the age effect and the sex effect need to be interpreted. An individual's suicide behavior is then determined by that person's sex and age, but not by the specific age-sex combination.

The number of degrees of freedom for this constrained model is derived as follows. There are 34×8 independent cells, $18 \times (3 - 1)$ parameters γ_{mt} for the rows, and $T(J - 1) = 3 \times 8$ parameters for the columns; and the model is identified by constraining the row parameters $\hat{\pi}_{ikt}^{AC\bar{X}}$ by the matrix **V** described above. This gives 272 − (36 + 24) = 212*df.*

The constrained model has a fit of $G^2 = 1{,}136.6$, $df = 212$, and the difference between it and the unconstrained three-budget model is relatively slight, given the large sample size ($G^2 = 1{,}136.6 - 1{,}085.9 = 50.6$, $df = 212 - 186 = 26$). The parameter estimates in Table 3[3] are similar to the parameter estimates in Table 2, especially the column parameters. To simplify the interpretation of the row parameters $\hat{\pi}_{ikt}^{AC\bar{X}}$, we also give average age-effect parameter estimates $\hat{\pi}_{it}^{A\bar{X}}$ and average sex-effect parameter estimates $\hat{\pi}_{kt}^{C\bar{X}}$. (For a derivation of these parameter estimates, see the section "Multinomial logit constraints on row parameters," above.)

The average sex-effect parameter estimates show that the expected budgets for males consist far more than average (i.e., much more than the probabilities given by the average) of latent budgets 1 and 3, whereas the expected budgets for females consist far more than average of latent budget 2. The average age-effect parameter estimates show that the first age group has an extremely high tendency to use the third latent budget, the younger age groups tend to use the first latent budget, and the older groups are most likely to use the second latent budget.

This application shows how the factorial structure of the row categories may be used to constrain the row parameters. We conclude that by constraining the row parameters with a multinomial logit model, the interpretation has been simplified considerably. Although this type of application is possible, by using the methodology of loglinear modeling with latent variables, it has not received attention thus far in the literature.

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. For more details, see Junger (1990). Among other things, three crime measures were gathered from the police registration:

[3]We do not give the multinomial logit parameters γ_{mt}, because these parameters are much harder to interpret than the conditional probabilities that they yield (see (14)).

TABLE 3
Parameter Estimates for Constrained Model

Age	Males $\hat{\pi}_{ikt}^{AC\overline{X}}$			Females $\hat{\pi}_{ikt}^{AC\overline{X}}$			Average Age Effect $\hat{\pi}_{it}^{A\overline{X}}$		
	$t=1$	$t=2$	$t=3$	$t=1$	$t=2$	$t=3$	$t=1$	$t=2$	$t=3$
10–15	.023	.002	.975	.268	.732	.000	.070	.143	.787
15–20	.498	.015	.486	.522	.478	.000	.505	.144	.352
20–25	.642	.022	.336	.497	.503	.000	.606	.143	.251
25–30	.648	.023	.329	.483	.517	.000	.605	.151	.244
30–35	.614	.038	.347	.352	.648	.000	.547	.195	.258
35–40	.552	.036	.412	.339	.661	.000	.494	.205	.301
40–45	.473	.071	.456	.183	.817	.000	.391	.281	.328
45–50	.420	.081	.499	.149	.851	.000	.334	.326	.340
50–55	.354	.121	.525	.090	.910	.000	.237	.472	.291
55–60	.304	.165	.532	.059	.941	.000	.191	.522	.287
60–65	.217	.247	.536	.029	.971	.000	.129	.587	.285
65–70	.104	.359	.537	.010	.990	.000	.060	.652	.288
70–75	.054	.417	.530	.004	.996	.000	.032	.670	.298
75–80	.017	.434	.549	.001	.999	.000	.009	.701	.289
80–85	.000	.518	.482	.000	1.000	.000	.000	.748	.252
85–90	.000	.442	.558	.000	1.000	.000	.000	.662	.338
90+	.005	.395	.600	.000	1.000	.000	.003	.598	.399

Cause of Death[a]	Column Parameters $\hat{\pi}_{jt}^{\overline{B}X}$			
	$\hat{\pi}_{+j}$	$t=1$	$t=2$	$t=3$
1	.330	.543	.416	.000
2	.005	.011	.003	.000
3	.040	.130	.001	.006
4	.383	.000	.340	.823
5	.050	.021	.098	.012
6	.059	.094	.001	.103
7	.018	.006	.018	.029
8	.053	.054	.082	.014
9	.062	.141	.041	.013
Total	1.000	1.000	1.000	1.000

	Average Sex Effect $\hat{\pi}_{kt}^{C\overline{X}}$		
	$t=1$	$t=2$	$t=3$
Males	.383	.151	.466
Females	.143	.857	.000
Average	.296	.408	.296

[a]See note to Table 1.

TABLE 4
Crime Among Four Ethnic Groups, by Age

Ethnicity	Age	Crime Pattern BDE[a] 000	100	010	110	001	101	001	111	Total
Moroccans	12–13	65	13	1	1	0	1	0	1	82
	14–15	43	12	0	4	2	2	0	2	65
	16–17	26	18	1	2	0	3	0	1	51
Turks	12–13	52	4	0	0	1	0	0	0	57
	14–15	73	16	1	2	3	0	0	0	95
	16–17	32	13	1	3	0	1	0	1	51
Surinamese	12–13	71	9	1	0	2	2	0	0	85
	14–15	54	10	0	1	0	1	1	1	68
	16–17	36	12	1	1	0	2	0	1	53
Dutch	12–13	78	5	0	0	2	3	0	0	88
	14–15	70	5	1	1	0	1	0	1	79
	16–17	26	3	1	1	3	1	0	2	37
Total		626	120	8	16	13	17	1	10	811

Note: Type of crime: B = property crime, D = aggression against persons, E = vandalism.

[a]0 = not registered, 1 = registered.

property crime, aggression against persons, and vandalism. The age of the youngsters was coded into three categories: 12–13, 14–15, and 16–17. This led to the data in Table 4. The relevant questions are (1) Are there one or more underlying crime measures? (2) Are these underlying crime measures the same for each of the ethnic groups? (3) If so, how is group membership (i.e., ethnic-group and age-group membership) related to these underlying crime measures?

Since the table is sparse, the power of tests will be relatively low. Therefore, we will not use the test results as evidence for acceptance of models, but rather as evidence for not rejecting them.

The first two questions can be answered readily by a constrained form of simultaneous latent class analysis similar to (21). For the explanatory variables, let age be denoted by A, indexed by i, and ethnic group by C, indexed by k. For the response variables, let the three crime measures be denoted by B, D, and E, indexed by j, m, and n, respectively. Then the latent budget model is

$$\frac{\pi_{ikjmn}}{\pi_{ik+++}} = \sum_{t=1}^{T} \pi_{ikt}^{AC\overline{X}} \pi_{jmnt}^{\overline{BDE}X}, \tag{23}$$

with the parameters $\pi_{jmnt}^{\overline{BDE}X}$ constrained by a multinomial logit model (18) with fixed scores w_{jmnh} and parameters ψ_{ht}. The fixed scores w_{jmnh} can be collected in an 8×3 matrix $\mathbf{W}$, where 8 is the number of response combinations defined by j, m, and n, and 3 is the number of variables. The first column contrasts the two categories of variable B, the second column contrasts the two categories of variable D, and the third column contrasts the two categories of variable E. Thus, the three variables B, D, and E are unrelated in $\mathbf{W}$, and the variables B, D, and E are independent at each level of the latent variable X. Therefore, $\pi_{jmnt}^{\overline{BDE}X}$ can be rewritten as $\pi_{jmnt}^{\overline{BDE}X} = \pi_{jt}^{\overline{B}X}\pi_{mt}^{\overline{D}X}\pi_{nt}^{\overline{E}X}$ (compare (19) with (21)). The parameters ψ_{ht} are collected in a $3 \times T$ matrix, where T is the number of latent budgets. There are thus $3 \times T$ parameters to be estimated.

Model (23) with the above constraints corresponds to the following loglinear model for the latent probabilities π_{ikjmn}:

$$\pi_{ikjnmt} = \lambda + (\lambda_i^A + \lambda_k^C + \lambda_{ik}^{AC}) + (\lambda_j^B + \lambda_m^D + \lambda_n^E) + \lambda_t^X + (\lambda_{it}^{AX} + \lambda_{kt}^{CX} + \lambda_{ikt}^{ACX}) + (\lambda_{jt}^{BX} + \lambda_{mt}^{DX} + \lambda_{nt}^{EX}). \tag{24}$$

Therefore, because of the multinomial logit model for the parameters $\pi_{jmnt}^{\overline{BDE}X}$, there are no direct relations between the response variables B, D, and E.

Model (23) with two latent budgets, or classes, and constraint $\pi_{jmnt}^{\overline{BDE}X} = \pi_{jt}^{\overline{B}X}\pi_{mt}^{\overline{D}X}\pi_{nt}^{\overline{E}X}$ has an adequate fit: $G^2 = 65.93$, df is 66 (see Table 5). Since the Pearson chi-square statistic is better approximated by the chi-square distribution in case of small frequencies, we also give it here: $X = 72.14$. The parameters $\hat{\pi}_{ikt}^{AC\overline{X}}$, $\hat{\pi}_{jt}^{\overline{B}X}$, $\hat{\pi}_{mt}^{\overline{D}X}$, and $\hat{\pi}_{nt}^{\overline{E}X}$ are given in the first columns (model 1) of Table 6. In the first latent

TABLE 5
Fit of Models for Table 4

	G^2	X^2	df
Model 1, No constraints	65.93	72.15	66
Model 2, $\lambda_{ikt}^{ACX} = 0$	70.30	80.74	72
Model 3, $\lambda_{ikt}^{ACX} = 0$, $\lambda_{kt}^{CX} = 0$	104.87	131.38	74
Model 4, $\lambda_{ikt}^{ACX} = 0$, $\lambda_{it}^{AX} = 0$	86.70	85.60	75
Model 5, $\lambda_{ikt}^{ACX} = 0$, λ_{it}^{AX} is linear in age	70.31	80.81	73

Note: Models are defined in terms of constraints that they impose upon the conditional row probabilities $\pi_{ikt}^{AC\overline{X}}$.

TABLE 6a
Latent Budget Estimates for Table 4
(Parameter Estimates $\hat{\pi}_{ikt}^{AC\bar{X}}$)

		Model 1		Model 2		Model 5	
Ethnicity	Age	$t = 1$	$t = 2$	$t = 1$	$t = 2$	$t = 1$	$t = 2$
Moroccans	12–13	.851	.149	.853	.147	.855	.145
	14–15	.693	.307	.703	.297	.701	.299
	16–17	.510	.490	.482	.519	.483	.517
Turks	12–13	1.000	.000	.935	.065	.935	.065
	14–15	.858	.142	.852	.148	.852	.148
	16–17	.649	.351	.694	.306	.696	.304
Surinamese	12–13	.921	.079	.930	.070	.931	.069
	14–15	.836	.164	.844	.156	.843	.157
	16–17	.714	.286	.680	.320	.682	.318
Dutch	12–13	.944	.056	.959	.041	.960	.040
	14–15	.929	.071	.906	.094	.905	.095
	16–17	.779	.222	.790	.210	.792	.208

budget, children have estimated probabilities of .063, .006, and .017 of being arrested for crimes B, D, and E, whereas in the second latent budget, children have estimated probabilities of .859, .219, and .213 of being arrested for those crimes. The first budget is thus a budget for children who have a very low probability of being arrested, whereas these probabilities are high for B and moderate for D and E in the second budget. The parameters $\pi_{ikt}^{AC\bar{X}}$ show the estimated probabilities of each of the latent budgets for each of the age-ethnicity groups. It shows that for each ethnic group, the probability of the (more criminal) second budget increases as they get older, that Moroccans have the highest probability of the second budget in each age group, and that the Turks show a rapid increase in the probability of the second budget as they get older. The Dutch have relatively low probabilities of the criminal budget, especially as they get older. This analysis appears to answer the first two questions: (1) The data provide no evidence against the existence of only two latent types of criminality, and (2) all ethnic groups have these in common. The model with two latent budgets gives an adequate fit.

The third question is also partly answered by the above analysis. However, we may wonder whether there is in fact any evidence of an ethnic-group effect, of an age-group effect, and of an interac-

TABLE 6b
Latent Budget Estimates for Table 4
(Parameter Estimates $\hat{\pi}_{jt}^{\bar{B}X}$, $\hat{\pi}_{mt}^{\bar{D}X}$, $\hat{\pi}_{nt}^{\bar{E}X}$)

Type of Crime[a]	Registration	Model 1 $t = 1$	Model 1 $t = 2$	Model 2 $t = 1$	Model 2 $t = 2$	Model 5 $t = 1$	Model 5 $t = 2$
B	0	.937	.141	.942	.135	.941	.135
	1	.063	.859	.058	.865	.059	.865
D	0	.994	.781	.993	.787	.993	.786
	1	.006	.219	.007	.213	.007	.214
E	0	.983	.787	.983	.794	.983	.794
	1	.017	.213	.017	.206	.017	.206

[a]See note to Table 4.

tion between age and ethnic group. This can be tested by constraining the parameters $\pi_{ikt}^{AC\bar{X}}$ by using multinomial logit models. In terms of the latent loglinear model, we can constrain the parameters λ_{it}^{AX},λ_{kt}^{CX}, and λ_{ikt}^{ACX} to be zero. Goodness-of-fit statistics for the models that we will discuss are given in Table 5.

Constraining the parameters λ_{ikt}^{ACX} to be zero implies that there are age-group and ethnic-group effects on the amount of criminality, but also that the age-group effect is the same for each ethnic group and that the ethnic-group effect is the same for each age group. This is Model 2 in Table 5. For this model the fit decreases to $G^2 = 70.30$, $df = 72$, $X^2 = 80.74$), which is still adequate. The difference between the fit of Model 2 and the fit of unconstrained Model 1 (see Table 5) is also not significant: $G^2 = 4.37$, $df = 6$. Therefore, there is insufficient evidence of any interaction between age and ethnic group in their relation to crime. The parameter estimates for Model 2 are given in the middle columns of Table 6. They differ only slightly from those for the unconstrained solution. A plot of the parameters $\hat{\pi}_{ik2}^{AC\bar{X}}$, showing the probabilities for both the constrained and the unconstrained models of the criminal latent budget, is given in Figure 3. The lines connecting the different age points for the four ethnic groups are slightly more regular than those in the constrained solution.

A more restrictive model constrains both λ_{it}^{AX} and λ_{ikt}^{ACX} to be zero. This model (Model 3 in Table 5) assumes no direct relationship between age and crime, but only a direct relationship between ethnic group and crime (specified by the parameters λ_{kt}^{CX}). The fit of this

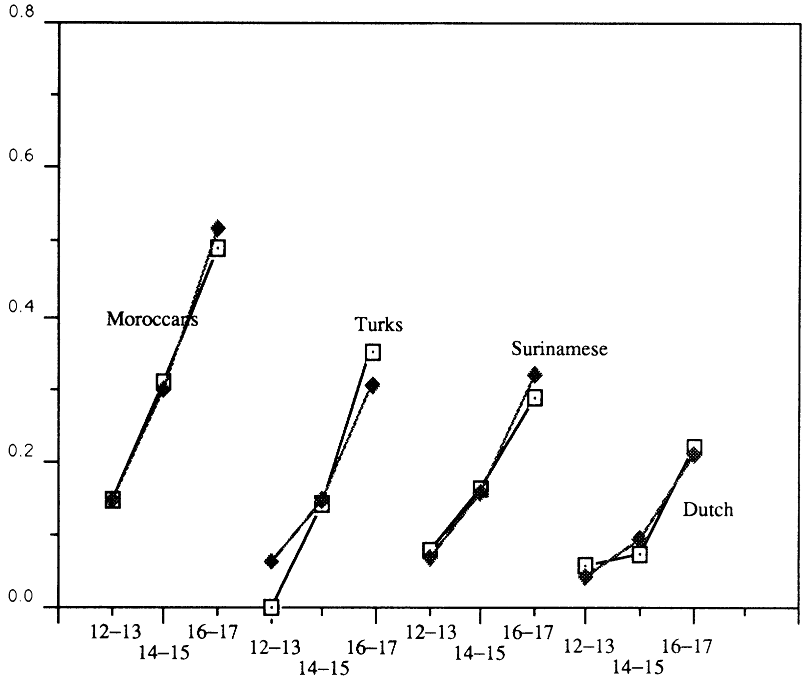

FIGURE 3. Plot of estimated row probabilities of the second (more criminal) latent budget. Open squares: unconstrained (Model 1, Table 4); solid squares: constrained (Model 2, Table 4).

model is poor: $G^2 = 104.87$, $df = 74$, $X^2 = 131.38$). Another model (Model 4) constrains both λ_{kt}^{CX} and λ_{ikt}^{ACX} to be zero. No ethnic-group effect is assumed, only an age-group effect. This model fits quite well, but we reject this hypothesis because the difference between it and Model 2 is significant: $G^2 = 86.70 - 70.30 = 16.40$, $df = 75 - 72 = 3$.

We therefore end up with Model 2. Model 2 can be constrained further by making use of the fact that the age categories are ordered. We can model this by replacing the two columns for age in the design matrix (one column for ages 12–13, and one for ages 14–15) with one column with values -1, 0, and 1 for ages 12–13, 14–15, and 16–17. Then, the age-group effect is assumed to be linear with age (Model 5). The difference between Models 2 and 5 is extremely small: $G^2 = 70.81$, $df = 73$, $X^2 = 80.81$.

The models used in this analysis are very similar to simultaneous latent class models. The example shows how multinomial logit constraints on the column parameters can be used to model conditional independence of three manifest variables given a latent variable. Assuming a latent variable that accounts for the association between manifest variables is standard in latent class analysis, but using the multinomial logit model for this purpose is new. The analysis shows that there is no evidence for more than two latent types of crime (the interpretation of the latent variable) and that all age-ethnic-group combinations have these two latent types of crime in common. Using the multinomial logit model to constrain the row parameters is also new. This turns out to be fruitful when there is additional information for the row variable. Here we can conclude that there is evidence of an age-group effect and of an ethnic-group effect, but not of age-ethnic-group interaction effects in the use of the criminal budgets. Furthermore, the age-group effect turns out to be linear in age.

4. CONCLUSION

Latent budget analysis is closely related to latent class analysis. In latent budget analysis the observed variables are considered to be either explanatory or response variables. In this sense latent budget analysis is different from latent class analysis, because in most applications of latent class analysis, all observed variables play the same role. However, the unconstrained models are equivalent, as shown in section 2. Latent class analysis of a two-way table is most often used when a latent variable should explain the association between the two manifest variables. Latent budget analysis can be used when the observed budgets are considered to be generated by a number of unknown latent budgets (see the mixture model in Figure 1) or when the latent variable is thought to intervene between two manifest variables (see the MIMIC model in Figure 2).

We have discussed constraints on the parameters of the latent budget model. The most interesting constraints discussed in the paper are the multinomial logit constraints. Our presentation emphasizes conditional independence of the rows and columns of a two-way table given a latent variable, and multinomial logit constraints imposed on the relations between the row variable and the latent variable and between the column variable and the latent variable. We showed that, just as for the unconstrained model, the constrained latent budget model can be understood as a loglinear model for the unobserved contingency table.

This conclusion suggests that constrained latent budget analysis is very similar to the loglinear model approach of latent class analysis as presented by, for example, Haberman (1979) and Hagenaars (1988, 1990). Indeed, although our models are different from the latent class models of Haberman and Hagenaars, our models can be fit using their methodology. However, the use of an explanatory variable for the rows and a response variable for the columns gives a specific interpretation to the parameters that is different from the interpretation in latent class analysis. In addition, the use of joint variables, and the exclusion of certain interaction effects, leads to interpretations that are different, and often more parsimonious, than those in latent class analysis. For example, Hagenaars (1988, 1990) discusses local dependence models as models with direct effects between manifest response variables and allows for these direct effects to account for correlated response error. In this paper direct effects between explanatory variables are included in the latent budget model because a multinomial distribution is assumed for each of the joint levels of the joint row variable. Both approaches can lead to the same loglinear model with latent variables, but with a different interpretation. The models given here allow us to answer questions previously unanswered by latent class models.

APPENDIX: SOFTWARE

The analysis in this paper has been performed with prototypes of programs written in APL. These prototypes can be obtained from the first author free of charge. However, it is also possible to use existing software.

Unconstrained latent budget analysis can be performed with existing programs for latent class analysis, such as MLLSA (Clogg 1977) or LCAG (Hagenaars and Luijkx 1990) or NEWTON, a program that replaces the earlier program LAT (see Haberman 1988). MLLSA and LCAG use the EM algorithm, while NEWTON uses a combination of the Newton-Raphson algorithm and the EM algorithm. After convergence of these programs, the row-parameter estimates of the latent budget model have to be derived using equation

(3). It is also possible to use the relation of latent budget analysis to simultaneous latent class analysis (see van der Heijden et al. 1989) or to mixed Markov latent class models (see van de Pol and Langeheine 1990). In programs that use the EM algorithm, simultaneous latent class models are fitted by using latent class models with so-called quasi-latent variables (see, for example, Hagenaars 1988, 1990). Then the row-parameter estimates are found directly. The program PANMARK (van de Pol, Langeheine, and de Jong 1989), used for mixed Markov latent class analysis, can be used if the latent budget model is considered as a mixed Markov latent class model for one variable with only one time point (see van de Pol and Langeheine 1990).

The above programs all allow for fixed-value and equality constraints in the latent budget model. Multinomial logit constraints can be fitted in NEWTON and LCAG by considering the latent budget model with multinomial logit constraints as a loglinear model with a latent variable. In NEWTON, design matrices are used to define the loglinear model for the unobserved matrix. Thus, all possible multinomial logit constraints can be fitted in NEWTON. In LCAG quasi-latent variables are used, and the program fits standard hierarchical loglinear models to the matrix with latent class probabilities (see Hagenaars 1988, 1990). Thus, all multinomial logit constraints that constrain loglinear interaction parameters to be zero can be fitted.

REFERENCES

Agresti, A. 1990. *Categorical Data Analysis.* New York: Wiley.

Aitchison, J. 1986. *The Statistical Analysis of Compositional Data.* London: Chapman and Hall.

Aitkin, M., D. Anderson, and J. Hinde. 1981. "Statistical Modeling of Data on Teaching Styles." *Journal of the Royal Statistical Society,* ser A., 144:419–61.

Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland. 1975. *Discrete Multivariate Analysis.* Cambridge, MA: MIT Press.

Bock, R. D. 1975. *Multivariate Statistical Methods in Behavioral Research.* New York: McGraw-Hill.

Clogg, C. C. 1977. "Unrestricted and Restricted Maximum Likelihood Latent Structure Analysis: A Manual for Users." Working Paper No. 1977–9. State College, PA: Pennsylvania State University, Population Issues Research Office.

———. 1981. "Latent Structure Models of Mobility." *American Journal of Sociology* 86:836–68.

Clogg, C. C., and L. A. Goodman. 1984. "Latent Structure Analysis of a Set of Multidimensional Contingency Tables." *Journal of the American Statistical Association* 79:762–71.

de Leeuw, J., and P. G. M. van der Heijden. 1988. "The Analysis of Time Budgets with a Latent Time-Budget Model." Pp. 159–66 in *Data Analysis and Informatics 5*, edited by E. Diday et al. Amsterdam: North Holland.

———. 1991. "Reduced Rank Models for Contingency Tables." *Biometrika* 78:229–32.

de Leeuw, J., P. G. M. van der Heijden, and P. Verboon. 1990. "A Latent Time-Budget Model." *Statistica Neerlandica* 44:1–22.

Dempster, A. P., N. M. Laird, and D. B. Rubin. 1977. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society,* ser. B, 39:1–38.

Everitt, B. S. 1988. "A Monte Carlo Investigation of the Likelihood Ratio Test for Number of Classes in Latent Class Analysis." *Multivariate Behavioral Research* 23:531–38.

Fienberg, S. E. 1980. *The Analysis of Cross-Classified Categorical Data.* 2d ed. Cambridge, MA: MIT Press.

Formann, A. K. 1978. "A Note on Parameter Estimation for Lazarsfeld's Latent Class Analysis." *Psychometrika* 43:123–26.

———. 1982. "Linear Logistic Latent Class Analysis." *Biometrical Journal* 24:171–90.

———. 1985. "Constrained Latent Class Models: Theory and Applications." *British Journal of Mathematical and Statistical Psychology* 38:87–111.

———. 1989. "Constrained Latent Class Models: Some Further Applications." *British Journal of Mathematical and Statistical Psychology* 42:37–54.

Gilula, Z. 1979. "Singular Value Decomposition of Probability Matrices: Probabilistic Aspects of Latent Dichotomous Variables." *Biometrics* 66:339–44.

———. 1983. "Latent Conditional Independence in Two-Way Contingency Tables: A Diagnostic Approach." *British Journal of Mathematical and Statistical Psychology* 36:114–22.

———. 1984. "On Some Similarities Between Canonical Correlation Models and Latent Class Models for Two-Way Contingency Tables." *Biometrika* 71:523–29.

———. 1986. "Grouping and Association in Contingency Tables: An Exploratory Canonical Correlation Approach." *Journal of the American Statistical Association* 81:773–79.

Gilula, Z., and S. J. Haberman. 1986. "Canonical Analysis of Contingency Tables by Maximum Likelihood." *Journal of the American Statistical Association* 81:780–88.

———. 1988. "The Analysis of Multivariate Contingency Tables by Restricted Canonical and Restricted Association Models." *Journal of the American Statistical Association* 83:760–71.

Gilula, Z., and A. M. Krieger. 1989. "Collapsed Two-Way Contingency Tables and the Chi-Square Reduction Principle." *Journal of the Royal Statistical Society,* ser. B, 51:425–33.

Good, I. J. 1969. "Some Applications of the Singular Decomposition of a Matrix." *Technometrics* 11:823–31.

Goodman, L. A. 1971. "Partitioning of Chi-Square, Analysis of Marginal Contingency Tables, and Estimation of Expected Frequencies in Multidimensional Contingency Tables." *Journal of the American Statistical Association* 66:339–44.

———. 1974. "Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models." *Biometrika* 61:215–31.

———. 1981. "Criteria for Determining Whether Certain Categories in a Cross-Classification Table Should be Combined, with Special Reference to Occupational Categories in Occupational Mobility Tables." *American Journal of Sociology* 87:612–50.

———. 1985. "The Analysis of Cross-Classified Data Having Ordered and/or Unordered Categories: Association Models, Correlation Models, and Asymmetry Models for Contingency Tables With or Without Missing Entries." *Annals of Statistics* 13:10–69.

———. 1986. "Some Useful Extensions to the Usual Correspondence Analysis Approach and the Usual Loglinear Approach in the Analysis of Contingency Tables" (with comments). *International Statistical Review* 54:243–309.

———. 1987. "New Methods for Analyzing the Intrinsic Character of Qualitative Variables Using Cross-Classified Data." *American Journal of Sociology* 93:529–83.

Grover, R. 1987. "Estimation and Use of Standard Errors of Latent Class Model Parameters." *Journal of Marketing Research* 24:298–304.

Grover, R., and V. Srinivasan. 1987. "A Simultaneous Approach to Market Segmentation and to Market Structuring." *Journal of Marketing Research* 24:139–53.

Haberman, S. J. 1979. *Analysis of Qualitative Data.* 2 vols. New York: Academic Press.

———. 1988. "A Stabilized Newton-Raphson Algorithm for Log-Linear Models for Frequency Tables Derived by Indirect Observation." Pp. 193–212 in *Sociological Methodology 1988,* edited by C. C. Clogg. Washington, DC: American Sociological Association.

Hagenaars, J. A. 1986. "Symmetry, Quasi-Symmetry, and Marginal Homogeneity on the Latent Level." *Social Science Research* 15:241–55.

———. 1988. "Latent Structure Models with Direct Effects Between Indicators. Local Dependence Models." *Sociological Methods and Research* 16:379–405.

———. 1990. *Categorical Longitudinal Data. Log-Linear, Panel, Trend, and Cohort Analysis.* London: Sage.

Hagenaars, J. A., and R. Luijkx. 1990. "LCAG. A Program for Latent Class Models and Other Loglinear Models with Latent Variables With and Without Missing Data." Working Paper No. 17. Tilburg: Tilburg University, Department of Sociology.

Heudin, H. 1982. *Suicide in America.* New York: Norton.

Heuer, J. 1979. *Selbstmord bei Kinder und Jugendlichen* (Suicide of children and youth). Stuttgard, Germany: Ernst Klett Verlag.

Junger, M. 1990. *Delinquency and Ethnicity.* Deventer, The Netherlands: Kluwer.

Langeheine, R. 1989. "New Developments in Latent Class Theory." Pp. 77–108 in *Latent Trait and Latent Class Models,* edited by R. Langeheine and J. Rost. New York: Plenum.

Luijkx, R. 1987. "Loglinear Modeling with Latent Variables: The Case of Mobility Tables." In *Sociometrics Research,* vol. 2, edited by W. Saris and I. Gallhofer. London: MacMillan.

Marsden, P. V. 1985. "Latent Structure Models for Relationally Defined Social Classes." *American Journal of Sociology* 90:1002–21.

McCutcheon, A. L. 1987. *Latent Class Analysis.* No. 64 in *Quantitative Applications in the Social Sciences.* London: Sage.

Mooijaart, A. 1982. "Latent Structure Analysis for Categorical Variables." In *Systems Under Indirect Observation,* edited by K. G. Jöreskog and H. Wold. Amsterdam: North Holland.

Mooijaart, A., and P. G. M. van der Heijden. 1992. "The EM Algorithm for Latent Class Analysis with Constraints." *Psychometrika* (in press).

Shigemasu, K., and Sugiyama, N. 1989. "Latent Class Analysis of Choice Behavior." Paper presented at the Annual Meetings of the Psychometric Society, University of California at Los Angeles.

Takane, Y. 1987. "Analysis of Contingency Tables by Ideal Point Discriminant Analysis." *Psychometrika* 52:493–513.

van de Pol, F., and R. Langeheine. 1990. "Mixed Markov Latent Class Models: From Description Towards Explanation." Pp. 213–47 in *Sociological Methodology 1990,* edited by C. C. Clogg. Oxford: Basil Blackwell.

van de Pol, F., R. Langeheine, and W. de Jong. 1989. *PANMARK User Manual.* Voorburg: Netherlands Central Bureau of Statistics.

van der Heijden, P. G. M. 1991. "Three Approaches to Study the Departure from Quasi-Independence." *Statistics Applicata* (in press).

van der Heijden, P. G. M., and J. de Leeuw. 1985. "Correspondence Analysis Used Complementary to Loglinear Analysis." *Psychometrika* 50:429–47.

van der Heijden, P. G. M., A. Mooijaart, and J. de Leeuw. 1989. "Latent Budget Analysis. In *Statistical Modelling. Proceedings, Trento, 1989,* edited by A. Decarli et al. Berlin: Springer Verlag.

Whittaker, J. 1990. *Graphical Models in Applied Multivariate Statistics.* New York: Wiley.