Multi-sample latent logit models with polytomous effects variables

McCutcheon, A.L.

Publication date:
1993

Document Version
Publisher's PDF, also known as Version of record

Work and Organization Research Centre

[redacted]

Rui
- models
- public opinion
- statistical method
- data analysis

PAPER

Multi-Sample Latent Logit Models with Polytomous Effects Variables

Allan L. McCutcheon
University of Delaware

WORC PAPER 93.08.014/7

September 1993

WORC papers have not been subjected to formal review or approval. They are distributed in order to make the results of current research available to others, and to encourage discussions and suggestions.

DRAFT: Do not quote or cite without the author's permission.

ACKNOWLEDGMENT

This paper has been prepared for presentation at the Symposium on Latent Structure Models for the Analysis of Social Change, sponsored by the Work and Organization Research Centre, Tilburg University, Tilburg, The Netherlands, September 8-10, 1993.

Multi-Sample Latent Logit Models with Polytomous Effects Variables

Allan L. McCutcheon
University of Delaware

Often we wish to study phenomena that are not directly observable. This has given rise to a family of measurement models for estimating (unobserved) latent variables from a set of observed variables. When the observable indicator variables are categorically measured, and theory and data suggest that the latent object of study is categorical, the latent class model (LCM) is appropriate for characterizing the latent variable. Such models may be of use in the study of learning and mastery models in educational research (Rindskopf 1983, Dayton and Macready 1988a, 1988b), ideal types in social research (Hagenaars and Halman 1989), market segmentation in marketing research, as well as many other areas of research involving categorical latent variables. In this paper we will explore the use of the latent class model (LCM) to study trends in public opinion.

In recent years, researchers have also shown an interest in examining associations between latent varaibles and *non*-indicator observed variables. In particular, models similar to Jöreskog and Goldberger's (1969) multiple indicator-multiple cause (MIMIC) models are suitable for examining such associations. These new models, which involve latent categorical variables and categorically-scored observed variables, are often estimated within the framework of loglinear models.

The purpose of this paper is to extend prior developments concerning the measurement and analysis of categorical latent variables to situations which include both indicator and non-indicator observed variables that are collected from two or more mutually exclusive populations. Thus, these models can be used in comparative research in which data on identical indicator and non-indicator variables have been collected in multiple samples. As we will see, this model is a latent class model derived from Haberman's (1979) loglinear parameterization of the LCM. Furthermore, it extends the loglinear LCM to include linear-by-linear restrictions on polytomous non-indicator variables. This advance provides a more efficient test of linear associations between latent variables and polytomous non-indicator variables than has previously been possible.

Background

In his early writing on the topic of latent class models, Goodman (1974a, 1974b) introduced an important breakthrough by demonstrating that the LCM could be estimated using the EM algorithm (Dempster, Laird and Rubin 1977). He further discussed how loglinear models with latent variables, such as loglinear MIMIC models, could be estimated as latent class models with restrictions on the model parameters (see also Clogg 1981). These models were readily estimated by Clogg's widely available program MLLSA (Clogg 1979). Hagenaars (1990) extended Goodman's original work with loglinear models with latent variables in two significant ways. First, he introduced a modification to the EM algorithm to allow the estimation of non-saturated loglinear models with latent variables. This contribution allows the estimation of non-saturated MIMIC models with a categorical latent variable. Second, he introduced a modification allowing for local dependence among the indicator variables (Hagenaars 1988). These models can be readily estimated using his widely available program LCAG (Hagenaars and Luijkx 1990).

Haberman (1976; 1979, chap. 10) provided another important breakthrough by showing how the latent class model could be estimated as a loglinear model with restricted parameters. Haberman also showed how a non-indicator, grouping variable could be in included in the latent class model. These models can be readily estimated with Haberman's widely available programs LAT (1979) and NEWTON (1988).

Following these developments, this author presented a method for the estimation of latent logit models (i.e., logit models with unobserved response variables) which takes advantage of the linearities that are potentially present in models with multilevel (i.e., polytomous) effects variables (McCutcheon forthcoming). The approach is an extension and generalization of earlier work presented by Haberman (1974, 1979) and Clogg (1988); it also advances earlier work by Goodman (1974a, 1974b) and more recent work by Hagenaars (1988, 1990). This latent logit model is closely related to the linear-by-linear restricted model first presented by Haberman (1974). The present paper is an extension of latent logit models to the multiple sample case in which one or more of the non-indicator variables is polytomous. It also examines how the latent loglinear model can be specified in an effort to overcome two of the common criticisms of loglinear models: parameter interpretability and parameter inflation. Unlike the intuitive interpretation available for the conditional and latent class probabilities of Lazarsfeld's LCM, the loglinear LCM is estimated as a set of logs of odds ratios (lambdas). As we will see, however, many of the most common parameter restrictions used with the Lazarsfeld LCM have equivalent forms in the loglinear LCM. Thus, many of the usual restrictions on conditional and latent class restrictions are also available with the loglinear LCM.

We will also consider a second common criticism of loglinear models--the parameter inflation problem. Within the usual loglinear context, a variable with I levels requires the estimation of I-1 parameters for every effect that includes that variable. As long as all variables in the analysis are dichotomous, the number of parameter estimates remains modest. When one or more polytomous (multi-level) variables are included in the loglinear model, however, the number of required parameters can grow rapidly. Dayton and Macready (1988a, 1988b) have considered models such as concomitant-variable latent class models which include a continuous, non-indicator effects variable in the latent class model. Formann (1992) has also recently extended the linear logistic latent class model (Formann 1982, 1985) to include polytomous data. This paper generalizes the loglinear LCM to include polytomous non-indicator variables, with a special focus on the multi-sample latent logit MIMIC model.

In the following sections, we first briefly explore the LCM restrictions on the latent class and conditional probabilities for MIMIC model using Lazarsfeld's parameterization. We also briefly consider the multi-sample extension of this model. Next, we examine Haberman's loglinear parameterization of LCM with extension to the MIMIC model with one or more polytomous non-indicator variables (McCutcheon forthcoming). As we will note, this approach is similar to the linear-by-linear parameter restrictions discussed by Haberman (1974), as well as the row and column effects in Goodman's Model I association models (Goodman 1979b, Agresti 1985). Special attention will be focused on: 1) how restricting the loglinear lambda parameters results in commonly desired restrictions on conditional and latent class probabilities; and 2) the extension of the model to the multi-sample case. In the third section, we consider the loglinear LCM as a latent logit model, in which the multi-sample model can be presented as a logit model with a latent dependent (response) variable. Finally, we consider an empirical example in which the latent logit model is used to examine changes in the American public's attitudes toward legalized abortion for social reasons during the two decades in which such abortions have been legal.

The Latent Class Model

In two early papers, Lazarsfeld presented a latent class model that would "explain" the association among a set of categorically-scored indicator variables in a manner analogous to factor analysis which "explains" the association among a set of continuously-scored indicator variables.[1] Although conceptually analogous to factor analysis, latent class analysis presupposes a categorical latent variable (X_t) with a set of T mutually exclusive and exhaustive classes, rather than a continuous latent variable as in factor analysis. Lazarsfeld's model assigns a probability for assignment to each of the T classes (π^X_t), with the restriction that $\sum_t \pi_t^X = 1.0$. Each level of the indicator variables is also assigned a probability conditional on the class of the latent variable (e.g. $\pi_{it}^{\bar{A}X}$), with the restriction that $\sum_i \pi_{it}^{\bar{A}X} = 1.0$. Thus, Lazarsfeld's basic latent class model expresses the crosstabulation of observed and unobserved variables as a function of the latent class and conditional probabilities. For example, when there are three indicator variables (e.g., A_i, B_j, C_k), the model is

$$\pi_{ijkt}^{ABCX} = \pi_t^X \times \pi_{it}^{\bar{A}X} \times \pi_{jt}^{\bar{B}X} \times \pi_{kt}^{\bar{C}X} . \qquad (1)$$

Goodman (1974a, 1974b, 1979a) made an important breakthrough by showing that the parameters of the latent class could be reliably estimated using iterative proportional fitting, a variant of the EM algorithm (Dempster, Laird, and Rubin 1977). Clogg (1979) implemented this algorithm in his widely available program MLLSA.

In another early contribution to latent class analysis, Goodman noted that the latent class probabilities (π^{X}_{t}) could be restricted to estimate loglinear models with latent variables (see esp. Goodman, 1974a, 1256-1257). Clogg (1981) has shown how this approach allows for the estimation of multiple indicator, multiple cause (MIMIC) models for categorical data, similar to the models introduced by Jöreskog for continuous data (Jöreskog and Goldberger 1969, Jöreskog 1973). Since the latent class probabilities (π^{X}_{t}) in (1) can be restricted to include both non-indicator and latent variables, it is possible to restrict a model with one latent (e.g., Z_r) and two non-indicator (e.g., E_m, F_n) variables to a T=R x M x N latent class model

$$\pi^{ABCX}_{ijkt} = \pi^{X}_{t} \times \pi^{\bar{A}X}_{it} \times \pi^{\bar{B}X}_{jt} \times \pi^{\bar{C}X}_{kt} \times \pi^{\bar{E}X}_{mt} \times \pi^{\bar{F}X}_{nt} \times \pi^{\bar{Z}X}_{rt}, \qquad (2)$$

where the conditional probabilities relating the latent variable X_t to the latent variable Z_r ($\pi^{\bar{Z}X}_{rt}$)

and the "*quasi*-latent," non-indicator variables E_m and F_n ($\pi^{\bar{E}X}_{mt}$, $\pi^{\bar{F}X}_{nt}$) are all restricted to either

0.0 or 1.0. In a case with dichotomous latent and quasi-latent variables (i.e., R=M=N=2), we restrict the conditional probabilities, mapping the latent class probabilities for X_t (T=8) on to a latent crosstabulation of the observed variables (E and F) and a dichotomous latent variable Z (see Figure 1). As can seen in Figure 1, we must impose the restrictions $\pi^{\bar{E}X}_{11} = \pi^{\bar{E}X}_{12} = \pi^{\bar{E}X}_{13} = \pi^{\bar{E}X}_{14} = 1.0$, which implies $\pi_{21}^{\bar{E}X}=\pi_{22}^{\bar{E}X}=\pi_{23}^{\bar{E}X}=\pi_{24}^{\bar{E}X}=0.0$. Once we have obtained estimates for the latent class probabilities π^{X}_{t}, we can use these values to estimate the parameters of the saturated[2] latent loglinear model (EFZ), where E and F are observed, and Z is latent. This latent loglinear model is illustrated in Figure 2. Models such as these can be estimated using several widely-available programs including MLLSA (Clogg 1979), LCAG (Hagenaars and Luijkx 1970), LAT (Haberman 1979), and NEWTON (Haberman 1988).

In recent work, Hagenaars (1988, 1990) has extended the latent loglinear model to include non-saturated models with latent and observed non-indicator variables. Hagenaars' method modifies the EM algorithm to adjust the relevant marginals of the latent crosstabulation at the end of each M-step (Hagenaars 1990, pg. 124). This modified EM algorithm has been implemented in the widely available program LCAG (Hagenaars and Luijkx, 1990).

Hagenaars' approach represents an important advance over the earlier methodology for estimating loglinear models with latent variables. His modification of the EM algorithm enables the extension of latent loglinear and logit models to the estimation of hierarchical models in which higher order terms are restricted to zero. Consequently, the range of non-saturated, hierarchical loglinear models, which previously had been limited to manifest (observed) variables, now extends to models having a combination of manifest and latent variables. Thus, Hagenaars' approach allows for the estimation of categorical data models that are analogous to continuous data MIMIC models (illustrated in Figure 3).

Goodman's and Hagenaars' approaches, however, are best suited to instances in which all variables other than X_t are dichotomous; neither approach is able to exploit potential linearities in associations between latent response variables and non-indicator variables that are ordered polytomies. This problem has been addressed in several recent papers. Dayton and Macready (1988a, 1988b) have examined concomitant-latent class models in which the latent class probabilities (π^{X}_{t}) depend on some multi-level covariate(s). Formann (1992) has also recently extended the linear-logistic LCM (Formann 1982, 1985) to include polytomous data. In a recent paper (McCutcheon, forthcoming) I have examined latent loglinear LCMs which include linear-by-linear restrictions on the parameters of polytomous effects variables. The current work is the multi-sample extension of the earlier work; both are directly related to the linear-by-linear restrictions in loglinear models first presented by Haberman (1974), as well as recent work by Goodman (1979b, see also Agresti 1985) on Model I association models.

Before turning to the loglinear LCM, we note Clogg and Goodman's (1985, 1986, 1987) extension of the LCM to multi-sample analyses. Where responses to the same stimuli are obtained in two or more mutually exclusive populations, such as samples from two or more nations, states, regions, or points in time, one of the non-indicator variables in the LCM (e.g., E_m) can be a sample variable. Thus two or more samples can be compared as simultaneous LCMs. With these simultaneous latent class models it is possible to restrict the measurement portion of the LCM relating the indicator variables to the latent variable, but it is somewhat more difficult to restrict the structural portion of the model which relates the non-indicator variables to the latent variable.

The Loglinear Latent Class Model

As Haberman notes (1979), the basic loglinear parameterization of the latent class model is

$$\log \hat{f}_{ijkt} = \lambda + \lambda_t^X + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX}. \qquad (3)$$

where X_t is the latent variable with T classes (t=1, ..., T), and A_i, B_j, and C_k are observed indicator variables. As with conventional loglinear models, we must impose the restriction that the lambdas sum to zero

$$\sum_t \lambda_t^X = \sum_i \lambda_i^A = \sum_j \lambda_j^B = \sum_k \lambda_k^C = \sum_t \lambda_{it}^{AX} = \sum_t \lambda_{jt}^{BX} = \sum_t \lambda_{kt}^{CX} =$$
$$\sum_i \lambda_{it}^{AX} = \sum_j \lambda_{jt}^{BX} = \sum_k \lambda_{kt}^{CX} = \sum_{i,t} \lambda_{it}^{AX} = \sum_{j,t} \lambda_{jt}^{BX} = \sum_{k,t} \lambda_{kt}^{CX} = 0$$

As Haberman also notes, a non-indicator variable may be introduced into the loglinear latent class model

$$\log \hat{f}_{ijktq} = \mu + \lambda_t^X + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX} + \lambda_{tq}^{XU}, \qquad (4)$$

where $\mu = \lambda + \lambda^U_q$.

In a manner analogous to Goodman's, we can extend Haberman's latent loglinear model by reparameterizing the non-indicator variable (U) into two or more quasi-latent variables. For example, we can define Uq as the joint distribution of two observed, non-indicator variables Em and Fn, where Q = M x N, and (q=1, .., N, N+1, ..., Q) is mapped onto m,n as [(1,1), ..., (1,N), (2,1), ..., (M,N)]. Thus the model in (4) may be rewritten as the loglinear MIMIC model

$$\log \hat{f}_{ijktmn} = \mu + \lambda_t^x + \lambda_i^a + \lambda_j^b + \lambda_k^c + \lambda_{it}^{ax} + \lambda_{jt}^{bx} + \lambda_{kt}^{cx} + \lambda_{tm}^{xe} + \lambda_{tn}^{xf} + \lambda_{tmn}^{xef}. \qquad (5)$$

Using the model presented in (5), it is possible to impose restrictions on the lambda parameters that are identical to restrictions on the conditional probabilities of Lazarsfeld's parameterization, especially as these relate to the restrictions of interest for loglinear MIMIC models. As Hagenaars (1990) has noted, equality restrictions are often imposed to obtain equal "error rates" for each of the indicator variables with respect to the latent variable; this is equivalent to the restriction that the indicators have identical rates for false positives and false negatives. To test this hypothesis using Lazarsfeld's parameterization, we impose equality restrictions on the conditional probabilities such as

$$\pi_{11}^{\bar{A}X}=\pi_{22}^{\bar{A}X} \qquad \pi_{11}^{\bar{B}X}=\pi_{22}^{\bar{B}X} \qquad \pi_{11}^{\bar{C}X}=\pi_{22}^{\bar{C}X}. \tag{6}$$

(Mooijaart and van der Heijden (1992), however, have cautioned against such across-class equality restrictions when using the EM algorithm.) Using the equivalence noted by Haberman (1979, 551), when T=2 with dichotomous indicator variables

$$\pi_{11}^{\bar{A}X}=\frac{\exp(\lambda_1^A+\lambda_{11}^{AX})}{\exp(\lambda_1^A+\lambda_{11}^{AX})+\exp(-\lambda_1^A-\lambda_{11}^{AX})}, \tag{7}$$

it is clear that by substituting (7) into (6), and recalling the restrictions on the lambdas, the equal error rate restrictions of (6) are equivalent to imposing the restriction that $\lambda_i^A=\lambda_j^B=\lambda_k^C=0$. Thus, when we impose equal error rate restrictions on each of the indicator variables in the model in (5), it becomes

$$\hat{f}_{ijkmn}=\mu+\lambda^x+\lambda^{ax}+\lambda^{bx}+\lambda^{cx}+\lambda^{xe}+\lambda^{xf}+\lambda^{xef}. \tag{8}$$

It is also possible to test the "parallel indicator" hypothesis—that particular indicators are equally reliable. As Hagenaars notes (1990, 110), we can test the hypothesis that two indicator variables are parallel indicators by imposing equality constraints on their conditional probabilities, such as $\pi_{11}^{\bar{A}X}=\pi_{11}^{\bar{B}X}$ and $\pi_{12}^{\bar{A}X}=\pi_{12}^{\bar{B}X}$. In the loglinear LCM this is equivalent to imposing the equality constraints $\lambda_1^{A}=\lambda_1^{B}$ and $\lambda_{11}^{AX}=\lambda_{11}^{BX}$.

In a recent paper, I have shown that the model in (5) can be parameterized to include linear-by-linear restrictions on the association between polytomous, non-indicator variables and a latent variable(s) (McCutcheon forthcoming). When one or more of the non-indicator variables in (5) have three or more ordered categories, the latent loglinear model can be reparameterized in a manner that differs significantly from Goodman's and Hagenaars' approaches. With their approaches estimating the saturated model, when E, F, and X are polytomous, requires the estimation of (T-1)(M-1) parameters for the {XE} relationship, (T-1)(N-1) for the {XF} relationship, and (T-1)(M-1)(N-1) for the {XEF} relationship. In contrast, linear-by-linear restrictions (Haberman 1974) on these lambda parameters may be tested for the loglinear model in (5). In this case, the model in (5) may be estimated with three linear-restricted parameters (ϕ) instead of the (T-1)(MN-1) lambda parameters previously required for the {XE}, {XF}, and {XEF} associations. Specifically:

$$\log \hat{f}_{ijktmn} = \mu + \lambda_t^x + \lambda_i^a + \lambda_j^b + \lambda_k^c + \lambda_{it}^{ax} + \lambda_{jt}^{bx} + \lambda_{kt}^{cx}$$
$$+ \phi^{xe}(u_t - \bar{u})(v_m - \bar{v}) + \phi^{xf}(u_t - \bar{u})(w_n - \bar{w}) + \phi^{xef}(u_t - \bar{t})(v_m - \bar{v})(w_n - \bar{w}), \tag{9}$$

where v_m is the score assigned to category *m* of the ordered polytomous variable E, u_t is the score assigned to class *t* of the latent variable X, and w_n is the score assigned to category *n* of the non-indicator variable F. Thus, this approach avoids the parameter inflation problem by requiring the estimation of fewer model parameters; these linear-by-linear restrictions allow us to incorporate into our estimation the information that may be inherent in the ordering of the polytomous variables.

The latent loglinear model in (9) is analogous to several loglinear models considered by Agresti (1984, see esp. chap 5). As Agresti notes, when model (9) does not fit the data well, but model (5) does, there may be intermediate models that are simpler than (5). For instance, if the relationship between E and X is not linear, $\phi^{XE}(u_t - \bar{u})(v_m - \bar{v})$ may be replaced by either of the more general terms λ_{tm}^{XE} or $\tau_t^{XE}(v_m - \bar{v})$, where the τ_t^{XE} effects reflect the deviation in the *t* levels of log($\hat{f}_{tm}$) from independence as a linear function of E, with slope τ_t^{XE} and $\sum_t \tau_t^{XE} = 0$. If E and F are ordered polytomies and X is an unordered polytomy,

$\phi^{XE}(u_t-\bar{u})(v_m-\bar{v})$ and $\phi^{XEF}(u_t-\bar{u})(v_m-\bar{v})(w_n-\bar{w})$ would be replaced by $\tau_t^{XE}(v_m-\bar{v})$ and $\tau_t^{XEF}(v_m-\bar{v})(w_n-\bar{w})$.

The Multi-sample Latent Logit MIMIC Model

Latent loglinear models may also be used to analyze data from several samples simultaneously. This is done by constraining some or all parameters to be equal over the groups, and testing for model invariance over the samples. Examples of multiple group analysis for usual latent class models have been discussed by Clogg and Goodman (1984, 1985, 1986). The groups may be different nations, states, regions, cultural subgroups, or--to examine social change--separate samples drawn from the same population at two or more time points (see e.g., McCutcheon 1987); indeed, the groups may be any mutually exclusive set of observations on which identical variables are measured.

In this section we demonstrate first how the measurement and structural sets of model coefficients can be used to test the equivalence of models in two or more populations. Later in this section, we consider the multi-sample latent loglinear model as a multi-sample latent logit model with manifest effects variables and a latent response variable. As we will see, the coefficients of the latent logit model are readily divided into sets that are analogous to the measurement model coefficients (e.g., λ^{ax}) and the structural model coefficients (e.g., λ^{xe}) in the continuous data MIMIC model; thus, we will refer to this model as the multi-sample latent logit MIMIC model.

We begin by considering the multi-sample extension of the model in (5). Consider a sample variable G with S (s=1,...,S) mutually exclusive sets of observations on the manifest variables A_i, B_j, C_k, E_m, and F_n. The multi-sample extension of (5) may be written as

$$\log \hat{f}_{ijktmns} = \mu + \lambda_t^X + \lambda_{ts}^{XG} + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{is}^{AG} + \lambda_{js}^{BG} + \lambda_{ks}^{CG} + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX} + \lambda_{its}^{AXG} + \lambda_{jts}^{BXG} + \lambda_{kts}^{CXG} + \lambda_{tm}^{XE} + \lambda_{tn}^{XF} + \lambda_{tmn}^{XEF} + \lambda_{tms}^{XEG} + \lambda_{tns}^{XFG} + \lambda_{tmns}^{XEFG}. \quad (10)$$

Here we see that the sample variable (G_s) is allowed to have associations with each of the variables in the model. Since our initial goal is usually to establish that the latent variable is identical (or at least similar) in all of the samples, our first step is usually to test the "across-sample parallel indicators" hypothesis. This step is equivalent to testing the factor invariance hypothesis in the continuous data MIMIC model (Marsh and Hocevar, 1985; Byrne et al., 1989). To test the hypothesis that the indicator variables (A_i, B_j, C_k) are parallel indicators over the *S* samples, we impose the restrictions

$$\lambda_{is}^{AG} = \lambda_{js}^{BG} = \lambda_{ks}^{CG} = \lambda_{its}^{AXG} = \lambda_{jts}^{BXG} = \lambda_{kts}^{CXG} = 0.$$

When the across-sample parallel indicators hypothesis can be accepted for most, or all, of the indicator variables, we can accept that the latent variable is identical in the *S* samples. In this case (10) reduces to a variant of (5)

$$\log \hat{f}_{ijktmns} = \mu + \lambda_t^X + \lambda_{ts}^{XG} + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX} + \lambda_{tm}^{XE} + \lambda_{tn}^{XF} + \lambda_{tmn}^{XEF} + \lambda_{tms}^{XEG} + \lambda_{tns}^{XFG} + \lambda_{tmns}^{XEFG}. \quad (11)$$

In addition, when all, or some, of the indicator variables can be characterized as parallel indicators over the S samples, we may be interested in testing the hypothesis of "equal error rates," which states that the likelihood of a "false positive" for an indicator variable is equal to the likelihood of a "false negative" for the same variable. To test the "equal error rate" hypothesis, we impose the restrictions that $\lambda_i^A = \lambda_j^B = \lambda_k^C = 0$. When both the across-sample parallel indicator and equal error rate hypotheses can be accepted, the model in (10) reduces to a variant of (8)

$$\log \hat{f}_{ijktmns} = \mu + \lambda_t^X + \lambda_{ts}^{XG} + \lambda_{it}^{AX} + \lambda_{jt}^{BX} + \lambda_{kt}^{CX} + \lambda_{tm}^{XE} + \lambda_{tn}^{XF} + \lambda_{tmn}^{XEF} + \lambda_{tms}^{XEG} + \lambda_{tns}^{XFG} + \lambda_{tmns}^{XEFG}. \quad (12)$$

We can, of course, also test the usual parallel indicator hypothesis that $\lambda_{it}^{AX} = \lambda_{jt}^{BX} = \lambda_{kt}^{CX}$.

Once we have established that the latent variable is (reasonably) equivalent in the S samples, we may also examine models (10) - (12) for *structural equivalence* in the parameters relating the manifest, non-indicator variables to the latent variable.[3] Specifically, to test the hypothesis that E and X are equivalently related in the S samples, we restrict $\lambda_{tms}^{XEG} = 0$. Similar hypotheses may be tested by imposing restrictions on each of the remaining structural parameters that involve the sample variable (i.e., λ_{tns}^{XFG}, λ_{tmns}^{XEFG}).

Finally, consider the case where the latent variable X is a dichotomous response (dependent) variable. Following the argument based on Bishop (1969), the logit model may be derived by subtracting the latent loglinear model for $\hat{f}_{ijk2mns}$ from the model for $\hat{f}_{ijk1mns}$. Thus, for (5) we obtain

$$\log\left(\frac{f_{ijk1mn}}{f_{ijk2mn}}\right)=\beta_0+\beta_i^A+\beta_j^B+\beta_k^C+\beta_m^E+\beta_n^F+\beta_{mn}^{EF}, \tag{13}$$

where the β coefficients equal twice the respective lambda coefficients (e.g., $\beta_0=2\lambda^X_1$, $\beta^A_i=2\lambda^{AX}_{i1}$, $\beta^E_m=2\lambda^{XE}_{1m}$) and each of the variables is effects-coded. As we see, M-1 logit parameters must be estimated for the {XE} association, N-1 logit parameters must be estimated for the {XF} association, and (M-1)(N-1) logit parameters must be estimated for the {XEF} association; thus, a total of (MN-1) must be estimated for these associations in the latent logit model presented in (13). For (9), on the other hand, we obtain

$$\log\left(\frac{\hat{f}_{ijk1mns}}{\hat{f}_{ijk2mns}}\right)=\beta_0+\beta_i^A+\beta_j^B+\beta_k^C +\beta_m^E(v_m-\bar{v})+\beta_n^F(w_n-\bar{w})+\beta_{mn}^{EF}(v_m-\bar{v})(w_n-\bar{w}), \tag{14}$$

where β_0, β^A_i, β^B_j, and β^C_k are defined as in (13), and β^E, β^F, and β^{EF} are equal to twice their respective ϕ (or τ) coefficients. When {XE}, {XF}, and {XEF} are linear-by-linear associations (i.e., when they are calculated from the ϕ's), it is necessary to estimate only a single parameter for each of these associations; thus, MN-4 fewer estimated parameters are required than in equation (13). When either E or F is an unordered polytomy, the respective β's are calculated from the estimated τ's; thus, the number of estimated parameters is greater than when all associations are linear-by-linear. For example, if E is an ordered and F is an unordered polytomy, then the {XF} and {XEF} each require the estimation of N-1 parameters. This model, then, would require the estimation of 2N-1 parameters, and would result in the estimation of MN-2N-2 fewer parameters than in (13).

The multi-sample latent logit MIMIC model with linear-by-linear restrictions on the polytomous effects variables can also be expressed in terms similar to that of (14). Thus, the model in (10) becomes

$$\log\left(\frac{\hat{f}_{ijk1mns}}{\hat{f}_{ijk2mns}}\right) = \beta_0 + \beta_s^G + \beta_i^A + \beta_j^B + \beta_k^C + \beta_{is}^{AG} + \beta_{js}^{BG} + \beta_{ks}^{CG} + \beta^E(v_m-\bar{v}) + \beta^F(w_n-\bar{w}) + \beta^{EF}(v_m-\bar{m})(w_n-\bar{n}) + \beta_s^{EG}(v_m-\bar{v}) + \beta_s^{FG}(w_n-\bar{w}) + \beta_s^{EFG}(v_m-\bar{v})(w_n-\bar{w}) \quad . \qquad (15)$$

As with the model in (14), the β coefficients for the indicator variables in (15) are equal to twice the corresponding λ, and the β coefficients for the non-indicator variables are equal to twice the corresponding ϕ (or τ) parameter. While linear-by-linear restrictions may also be applied to polytomous sample (G_s), latent (X_t), and indicator (A_i, B_j, C_k) variables, only polytomous effects (non-indicator) variables are considered here.

An Empirical Example

The current example focuses on changes in Americans' attitudes toward legal abortions for social reasons over the past twenty years; this example extends the research on this topic that I have reported elsewhere (McCutcheon 1987b, forthcoming). The data analyzed here come from the 1972 and 1991 General Social Surveys (GSS). The GSS is conducted each spring (March and April) by the National Opinion Research Center (NORC) at the University of Chicago. The first of these two years (1972) includes data collected just prior to the 1973 U.S. Supreme Court decision in Roe v. Wade which legalized abortion during the first 6 months of pregnancy as long as the procedure was performed by qualified medical personel at appropriate medical facilities.

In the 1972 and 1991 GSS, respondents were asked a series of questions regarding their opinions about legal abortion. Three of these questions asked respondents about their approval (disapproval) of legalized abortion for women wanting an abortion for social (non-medical) reasons. Responses to these three questions serve as the indicator variables in the analyses reported here. Specifically, respondents were asked

> "Should it be possible for a pregnant woman to obtain a legal abortion if . . ."
> "If she is not married and does not want to marry the man?" (S_i)
> "If the family has a very low income and cannot afford any more children?" (P_j)
> "If she is married and does not want any more children?" (N_k)

Responses of "yes" or "no" were allowed for these questions.

In each of these two surveys, respondents were also asked about their attitudes toward premarital sex. One question asked:

> "If a man and woman have sex relations before marriage, do you believe that it is always wrong, almost always wrong, wrong only sometimes, or not at all wrong?" (M_n)

For the analysis presented here, the "almost always wrong" and "wrong only sometimes" responses were combined, yielding a three-level polytomous response variable. A second non-indicator variable, religion (R_m), was also included in the analysis; only Protestants and Catholics were included in the analysis, however. Finally, only white respondents were included in the current analysis.

Typically, the first issue that we should address in any analysis has to do with the appropriateness of the proposed analytic model. The categorically-scored responses for the indicator and non-indicator variables suggest that the LCM is appropriate; however, we must also consider the likely nature of the latent construct. As the bi-modal distribution in Figure 4 illustrates, American public opinion is highly polarized on the issue of legalized abortion for social reasons. In 1972, the year prior to the U.S. Supreme Court's controversial Roe v. Wade decision, nearly 80% ot the respondents reported either approval or disapproval of legalized abortion in *all three* scenarios depicted in the indicator questions; only one in five gave mixed responses. By 1991, the level of polarization appears to have risen, with 85% of the respondents indicating consistent approval or consistent disapproval. Consequently, it appears that for both samples, it is plausible to hypothesize a *categorical* latent variable in which respondents either approve or disapprove of legal abortions for social reasons. In this case, mixed responses are assumed to be erroneous.

The initial step is to test whether a T-class model fits the samples adequately. As the data reported in Table 1 indicate, the *saturated* latent loglinear model with a two class latent abortion variable fits the observed, multi-sample data reasonably well (L^2=70.87, 60 df, p>.1). Since this model is unrestricted, however, we can infer only that the two-class latent variable model is plausible for the two samples; there is no assurance, however, that the latent abortion attitude variable for 1972 is similar to the same variable in 1991.

Without some assurance that the dependent variable is the same in each of the samples, further analysis is of questionable value. Thus, the next concern in multi-sample analyses such as these is to establish construct invariance across the samples. Consequently we focus our attention on the measurement portion of the latent loglinear model: that is, on those lambda parameters that include both the indicator variables (i.e., S_i, P_j, N_k) and the sample variable (Y_s).

The first in our hierarchy of hypotheses (H_1) addresses the issue of whether there has been a significant change in the distribution of the indicator variables that is independent of the latent variable. Restricting the three lambda parameters (λ^{SY}, λ^{PY}, λ^{NY}) equal to zero nets a modest increase in the L^2 (73.72 - 70.87 = 2.85, 63 - 60 = 3 df, p>.2). Thus, these data indicate that between 1972 and 1991 any change in the distribution of the indicator variables is attributable to changes in the latent variable.

As with multi-sample linear structural equation models, the optimal case for multi-sample latent loglinear models is that in which there is complete invariance in the latent construct (see e.g., Marsh and Hocevar 1985, Jõreskog 1971). Thus, the second hypothesis (H_2) tests whether the indicator variables maintain a constant level of association with the latent variable across the two years. As the data in Table 1 indicate, restricting the three lambda parameters (λ^{SXY}, λ^{PXY}, λ^{NXY}) to equal zero results in an unacceptable erosion of the model L^2 (83.04 - 73.72 = 9.32, 66 - 63 = 3 df, p<.03). Thus, one or more of the indicator variables have experienced a significant change with respect to their association to the latent variable.

Although complete across-sample invariance in the latent construct is desired, it is not required (see e.g., Byrne et al. 1989). H_3 tests the partial invariance of the latent variable from 1972 to 1991. Here we allow the parameter for the "single woman" indicator (λ^{SXY}) to vary freely, resulting in an acceptable increase in the model L^2 (77.47 -73.72 = 3.75, 65 - 63 = 2 df, p>.1). From the negative sign of the parameter estimate reported in Table 2 it appears that the association of the "single woman" indicator was lower in the first sample (1972) than in the second (1991); other than the increase in the strength of the association between the latent variable and the "single woman" indicator, the latent variable for the 1972 sample is identical to the latent variable in 1991. This idicates a very high degree of across-sample invariance.

Once we have established the level of invariance in the measurement portion of the model, we turn to the final set of hypotheses regarding the measurement of the latent variable. In H_4 we test the "equal error" rate hypothesis by restricting the indicator lambdas (λ^S, λ^P, λ^N) to zero. As discussed earlier in (6) and (7), these restrictions can be interpreted as the "equal error rate" hypothesis in Lazarsfeld's parameterization of the LCM. There is another, equally interesting aspect of these restrictions: they test the hypothesis that the latent variable is the sole source of deviation from an equiprobable distribution for the indicator variables. As we can see from the results reported in Table 2, the unacceptably large increase in the model L^2 (97.77 - 77.47 = 20.30, 68 - 65 = 3, p<.001) indicates that we must reject H_4. In model H_5 which tests the equal error rate hypothesis for the "single woman" and "no more children" indicator variables only, we see that the hypothesis may be accepted (79.56 - 77.47 = 2.09, 67 - 65 = 2, p>.3); the association between the latent variable and legalized abortion for "poor women" can not be restricted in this manner. As the λ^P estimate of -.278 indicates, after accounting for the respondents' latent attitude toward social reasons for abortion, the log odds-ratio $2*\lambda^S = -.556$ and the odds-ratio is $e^{-.556} = .573$. Thus, after accounting for the latent attitude, the odds are estimated to be .573 that respondents have non-favorable attitude (disapproval) towards legal abortion for women who are married and feel they cannot afford anymore children.

With H_6 we shift the focus from the measurement model to the structural portion of the model. Specifically, we test the hypothesis that $\lambda^{XY}=0$: that between 1972 and 1991 there were no significant shifts in the distribution of the latent variable that are unaccounted for by the effects variables of religion and attitudes toward premarital sex. As the results in Table 1 indicate (79.60 - 79.56 = .04, 68 - 67 = 1, p>.75), we can accept this hypothesis. In H_7 we report tests of hypotheses that the higher order effects in the structural portion of the model can be restricted to zero. These hypotheses include tests of the effects of interactions among the independent variables on the latent variable (λ^{XMR}=0), as well as tests of hypotheses that the strength of the associations between the independent and latent variables have remained constant over the past 20 years.[4] As the results for H_7 indicate (reported in Table 1), we can accept the hypotheses that 1) the interaction between religion and premarital sex attitudes does not have a significant effect on the latent abortion attitude variable, and that 2) between 1972 and 1991 there were no significant shifts in the effects of religion and attitudes toward premarital sex on attitudes toward social reasons for abortion (88.07 -79.60 = 8.47, 75 - 68 = 7 df, p>.3).

The final hypothesis tests a linear-by-linear restriction on the association between the polytomous attitude toward premarital sex variable and the latent variable reflecting attitudes toward social reasons for abortion. As the results reported in H_8 indicate, this hypothesis must be rejected because it results in an unacceptably large erosion of the L^2 (94.44 - 88.07 = 6.37, 76 - 75 = 1 df, p<.025). Although we would normally reject the model of H_8 on empirical grounds, we will use this model to illustrate the interpretation of the model parameters.

Using (7) we can obtain the modeled conditional probabilities of approval for each of the indicator variables in the two latent classes in each of the two samples. As we see from the measurement model portion of the data reported in Table 3, the "poor women" indicator variable does not maintain an equal error rate over this time period; the significant λ^P parameter results in the probability of a "false negative" (.033) being substantially lower than the probability of a "false positive" (1.00 - .905 = .095). Although the remaining two indicator variables do maintain equal error rates in both samples, we note that the significant λ^{SXY} parameter results in a decrease in the error rate from .054 in 1972 to .021 in 1991.

We can also use the parameter estimates from the structural portion of the model to estimate the latent logit parameters (β), as well as the probability that a respondent will be at level 1 (class I) of the latent variable, given their position with respect to the effects variables. These probabilities are reported in Table 4. As these data show, at each level of the attitude toward premarital sex variable, Protestants are less likely than Catholics to hold disapproving attitudes toward legal abortion for social reasons. Only among those Protestants who believe that premarital sex is "not at all wrong," however, do we see a less than .50 likelihood of disapproval of legal abortion for social reasons. Finally we note from both Tables 2 and 4 that attitudes toward premarital sex have a substantially greater impact than does religion on the likelihood of disapproving (approving) of legal abortions for social reasons; among Catholics there is a difference in the likelihoods of more than .3 between the extremes of the premarital sex variable, and among Protestants there is a difference of nearly .4.

Summary and Conclusions

The advent of loglinear and logit modles has extended to the analysis of categorically-scored data much of the power and flexibility that was once available only through regression analysis of continuously-scored data. The application of regression models to categorical data, however, has long been know to result in mis-estimation (Aldrich and Nelson 1984, Hosmer and Lemeshow 1989, Nerlove and Press 1973). Consequently, logit models have played an especially important role in categorical data analysis, because logit models are most directly analogous to regression analysis.

Unlike the usual loglinear models, in which all of the variables have the same status, in logit models one of the variables is designated as the dependent (response) variable and the others are designated as causal (effects) variables (Agresti 1990, Aldrich and Nelson 1984, Bishop, Feinburg, and Holland 1975, Haberman 1979, Hagenaars 1990, Hosmer and Lemeshow 1989, Nerlove and Press 1973). Unlike regression analysis, however, logit models have discrete, categorically scored dependent variables. Consequently, researchers who wish to investigate causal models with categorically scored dependent variables often rely on logit analysis.[5]

The multi-sample latent logit model presented here, while closely related to those of Goodman and Hagenaars, derives directly from the latent class model first presented by Haberman (1979). Haberman examined the latent class model as a restricted loglinear model in which the indicator variables (e.g., A, B, C) are locally independent with respect to the latent variable (X). Thus, Haberman's basic model is analogous to Goodman's basic model. As I have shown elsewhere (McCutcheon forthcoming) Haberman's model can be extended to include observed variables (e.g., E_m, F_n) which are not indicator variables. The multi-sample latent logit model presented here illustrates that a non-indicator sample variable may also be included in the analysis.

Finally, when one or more of the non-indicator variables have three or more ordered categories, the multi-sample latent loglinear model can be reparameterized in a manner that differs significantly from Goodman's and Hagenaars' approaches. In the saturated model, for example, when M, N, and T are all greater than 2, the previous approaches require the estimation of (T-1)(M-1) lambda parameters for the {XE} relationship, (M-1)(N-1) lambda parameters for the {EF} relationship, and (T-1)(N-1)(M-1) for the {XEF} relationship. In contrast, single linear-by-linear restrictions on these lambda parameters may be tested for the loglinear model. This approach allows us to incorporate into our estimation the information that may be inherent in the ordering of the polytomous effects variables, thereby requiring the estimation of fewer model parameters and thus reducing the parameter inflation problem. Because the presence if linerities is directly tested, the approach is recommended for all instances involving ordered polytomies.

Notes

[1] Good introductions to the Lazarfeld's parameterization of the LCM can be found in McCutcheon (1987), Shockey (1988), and Langeheine (1988).

[2] Unlike saturated loglinear models, saturated latent loglinear models require certain parameters to be restricted to zero; the axiom of local independence allows indicator variables to have associations to have non-zero associations with only the latent variable(s).

[3] Marsh and Hocevar (1985) argue that there should be complete factoral invariance across the samples, though Byrne et. al (1989) make the case for partial invariance.

[4] Although reported here as a single 7 df test, each of the parameter constraints implied in H_7 were tested individually, and were acceptable at the .05 alpha level.

[5] Probit models provide an alternative to logit models (see e.g., Finney 1971). Since these models tend to yield similar results, the relative computational ease of the logit model and the greater ease in interpretation of the logit parameters have led most researchers to prefer the logit to the probit model.

References

Agresti, A. (1985) *Analysis of Ordinal Categorical Data*. New York: Wiley.

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

Aldrich, J. H. and F. D. Nelson (1984) *Linear Probability, Logit, and Probit Models*. Newbury Park, CA: Sage Publications.

Byrne, B. M., R. J. Shavelson, and B. Muthen (1989) "Testing for the equivalence of factor covariance and mean structures: The issue of partial invariance," *Psychological Bulletin*, **105**, 456-466.

Bishop, Y. V. V. (1969) ``Full contingency tables, logits, and split contingency tables," *Biometrics* **25**: 119-128.

Bishop, Y. V. V., S. E. Feinberg, and P. M. Holland (1975) *Discrete Multivariate Analysis: Theory and Practice*. Cambridge, MA: MIT Press.

Clogg, C. C. (1977) "Unrestricted and restricted maximum likelihood latent structure analysis: A manual for users." Working paper 1977-09. University Park, PA: Populations Issues Research Office.

Clogg, C. C. (1981) "New developments in latent structure analysis." In D. M. Jackson and E. F. Borgatta (eds.) *Factor Analysis and Measurement*. Beverley Hills, CA: Sage.

Clogg, C. C. (1988) "Latent Class Models for Measuring." In R. Langeheine and J. Rost (eds.) *Latent Trait and Latent Class Models*. New York: Plenum.

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**: 762-771.

Clogg, C. C. and L. A. Goodman (1985) "Simultaneous Latent Structure Analysis in Several Groups." In N. B. Tuma (ed.) *Sociological Methodology*. San Fransisco: Josey-Bass.

Clogg, C. C. and L. A. Goodman (1986) "On scaling models applied to data from several groups," *Psychometrika*, **51**: 123-135.

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

Dayton, C. M., and G. B. Macready (1988b) "A latent class covariate model with applications to criterion-referenced testing." In R. Langeheine and J. Rost (eds.) *Latent Trait and Latent Class Models*. New York: Plenum.

Dempster, A. P., N. M. Laird, and D. B. Rubin (1977) "Maximum likelihood from Incomplete data via the EM algorithm (with discussion)," *Journal of the Royal Statistical Society, series B* **39**: 1-38.

Finney, D. J. (1971) *Probit Analysis, 3rd edition*. Cambridge: Cambridge University Press.

Formann, Anton K. (1982) "Linear logistic latent class analysis," *Biometrical Journal*, **24** 171--190.

Formann, Anton K. (1985) *Die Latent-Class-Analyse: Einführung in Theorie und Anwendung*. Weinheim: Beltz.

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

Goodman, L. A. (1974a) "The analysis of systems of qualitative variables when some of the variables are unobservable. Part I -- A modified latent structure approach," *American Journal of Sociology*, **79**: 1197--1259.

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

Goodman, L. A. (1979a) "On the estimation of parameters in latent structure anlaysis," *Psychometrika* **44**: 123-128.

Goodman, L. A. (1979b) "Simple models for the analysis of association in cross-classifications having ordered categories," *Journal of the American Statistical Association* **74**: 537-552.

Haberman, S. J. (1974) Log-linear models for frequency tables with ordered classification," *Biometrics*, **30**: 589--600.

Haberman, S. J. (1979) *Analysis of Qualitative Data: Vol. 2 New Developments*. New York: Academic Press.

Hagenaars, J. A. (1988) "Latent structure models with direct effects between indicators: Local dependence models," *Sociological Methods and Research*, **16**: 379--405.

Hagenaars, J. A. (1990) *Categorical Longitudinal Data*. Newbury Park, CA: Sage.

Hagenaars, J. A. and L. C. Halman (1989) "Searching for ideal types: the potentialities of latent class analyses," *European Sociological Review* **5**: 81-96.

Hagenaars, J. A. and R. Luijkx (1990) "LCAG: A Program to estimate latent class models and other loglinear models with latent variables with and without missing data." Working Paper Series # 17^2. Tilburg, The Netherlands: Department of Sociology of Tilburg University.

Hosmer, D. W. and S. Lemeshow (1989) *Applied Logistic Regression*. New York: John Wiley.

Jöreskog, K. G. (1971) "Simultaneous factor analysis in several populations," *Psychometrika* 57: 409-426.

Jöreskog, K. G. (1973) "A general method for estimating a linear structural equation system." In A. S. Goldberger and O. D. Duncan (eds.) *Structural Equation Models in the Social Sciences*. New York: Seminar Press.

Jöreskog, K. G., and A. S. Goldberger (1975) "Estimation of a model with multiple indicators and multiple causes of a single latent variable," *Journal of the American Statistical Association* **10**: 631-639.

Langeheine, R. (1988) "New developments in latent class theory." In R. Langeheine and J. Rost (eds.) *Latent Trait and Latent Class Models*. New York: Plenum.

Marsh, H. W., and D. Hocevar (1985) "The application of confirmatory factor analysis to the study of self-concept: First and higher order structures and their invariance across age groups," *Psychological Bulletin*, **97**, 562--582.

McCutcheon, A. L. (1987a) *Latent Class Analysis*. Sage: Newbury Park, CA.

McCutcheon, A. L. (1987b) "Sexual morality, pro-life values, and attitudes toward abortion," *Sociological Methods and Research*, **16**: 256--275.

McCutcheon, A. L. (forthcoming) "Logit Model with Latent Dependent and Polytomous Response Variables." In A. von Eye and C. C. Clogg (eds.) *Analysis of Latent Variables in Developmental Research*. Newbury Park, CA: Sage.

Mooijaart, A., and P. G. M. van der Heijden (1992) "The EM algorithm for latent class models with equality constraints," *Psychometrika* **57**: 261-269.

Nerlove, M. and S. J. Press (1973) *Univariate and Multivariate Log-Linear and Logistic Models*. Rand Corp. Technical Report R-1306-EDA/NIH, Santa Monica, CA.

Rindskopf, D. (1983) "A general framework for using latent class analysis to test hierarchical and nonhierarchical learning models," *Psychometrika* **48**: 85-97.

Shockey, J. (1988) "Latent class analysis: An introduction to discrete data models with unobserved variables." In J. S. Long (ed.) *Common Problems/Proper Solutions*. Newbury Park, CA: Sage.

		Variable E: 1		Variable E: 2	
		Variable Z: 1	Variable Z: 2	Variable Z: 1	Variable Z: 2
Variable F	1	π^X_1	π^X_2	π^X_5	π^X_6
	2	π^X_3	π^X_4	π^X_7	π^X_8

Figure 1: Latent E × F × Z Crosstabulation

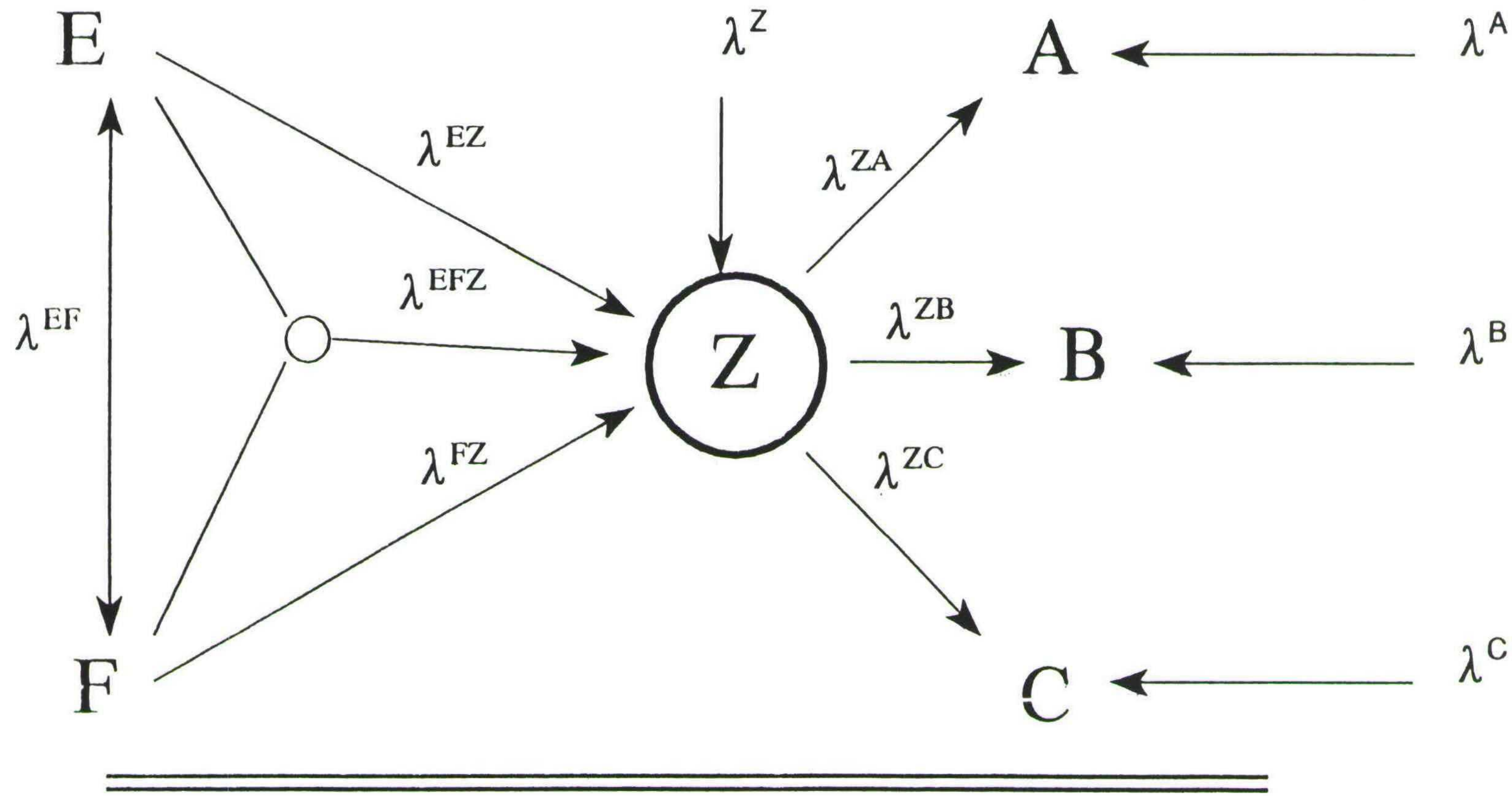

Figure 2: Saturated Loglinear Model with A Latent Dependent Variable (Z)

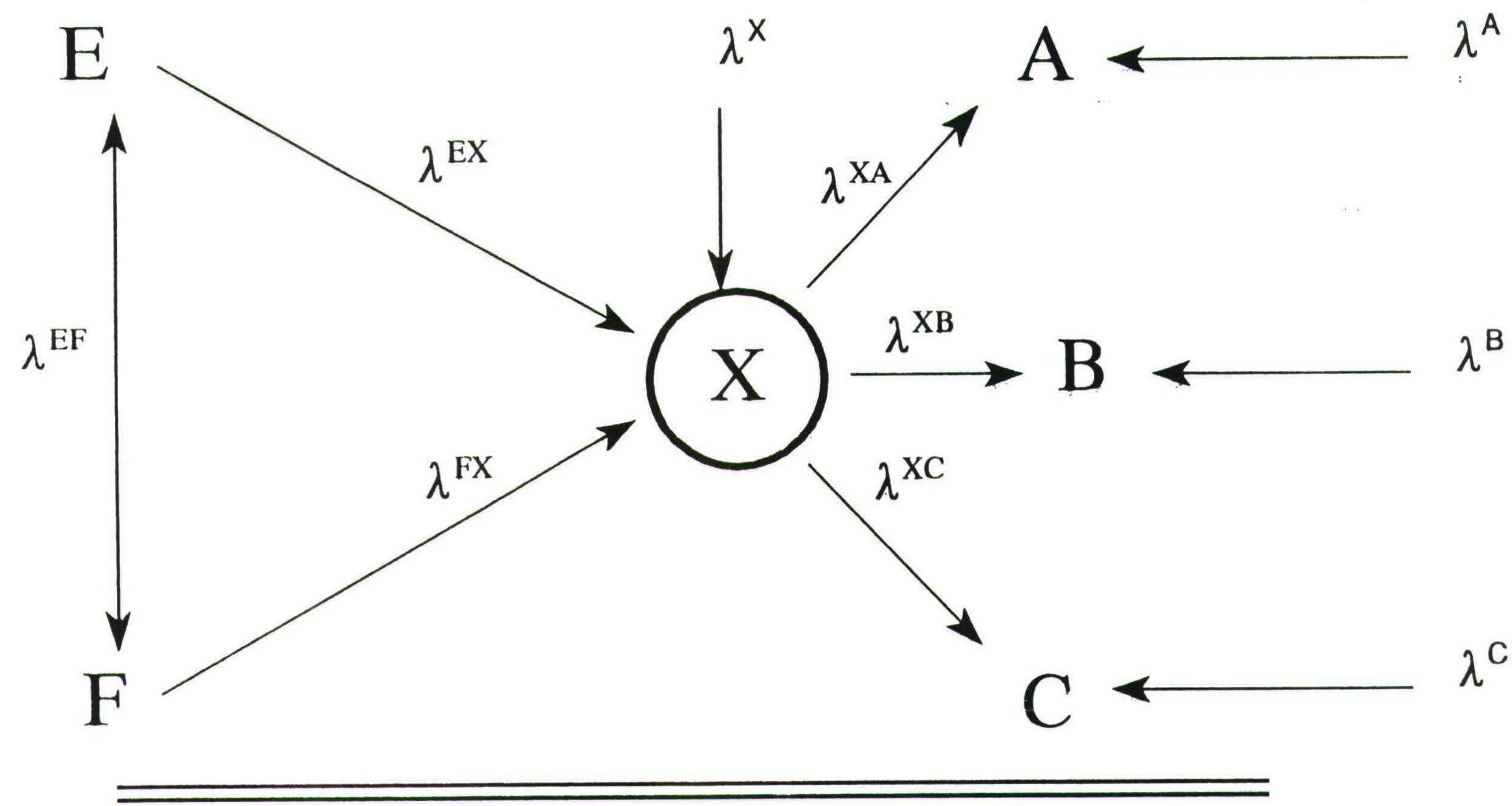

Figure 3: Loglinear Model with A Latent Dependent Variable (X)

Figure4: Approval of Social Reasons for Abortion

1972 - 1991

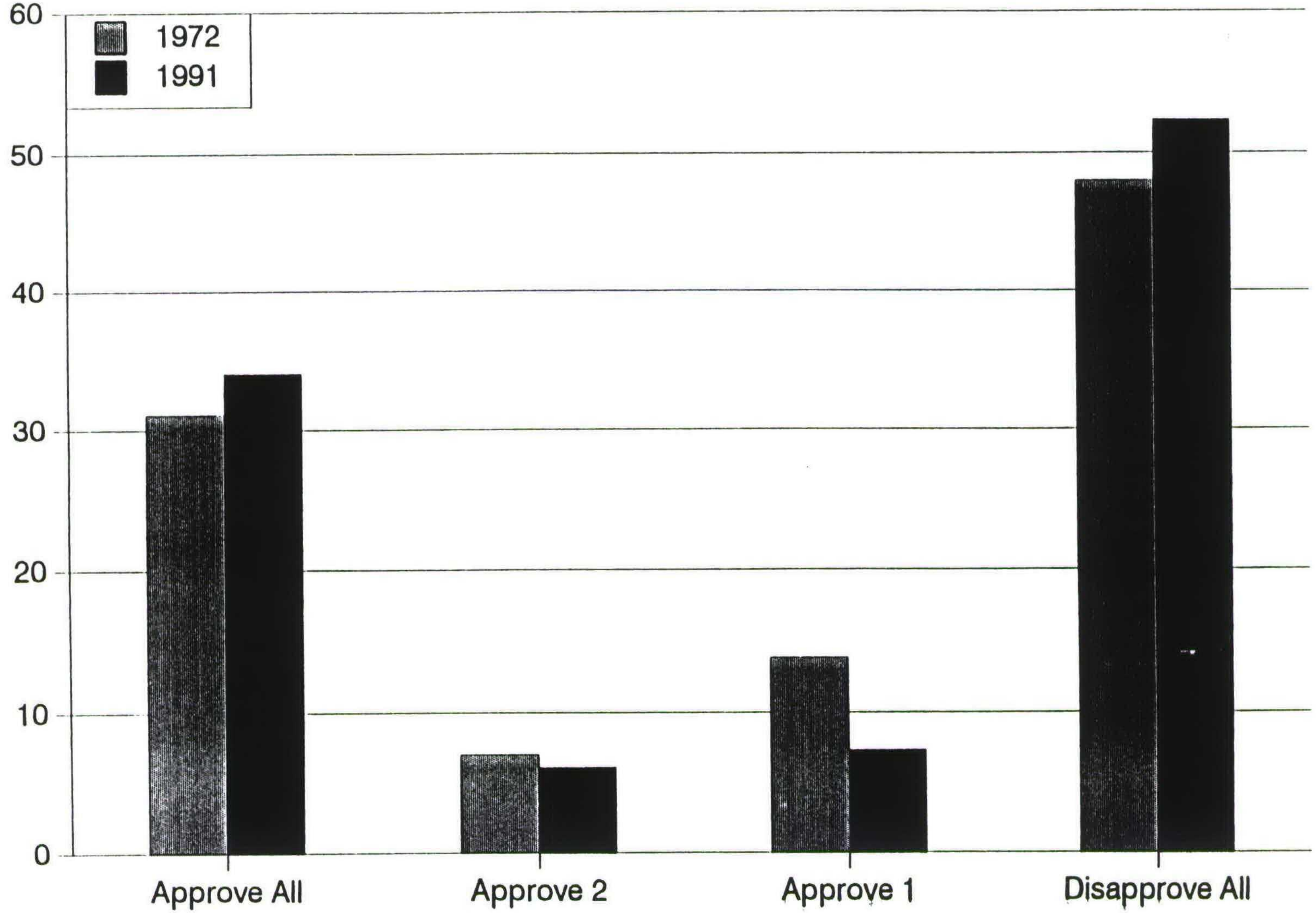

Table 1: Likelihood-ratio Chi-Squares for Selected Latent Logit MIMIC Models

Model	L^2	df
H_0: Saturated Latent Logit	70.87	60
H_1: $\lambda^{SY}=\lambda^{PY}=\lambda^{NY}=0$	73.72	63
H_2: $H_1 + \lambda^{SXY}=\lambda^{PXY}=\lambda^{NXY}=0$	83.04	66
H_3: $H_1 + \lambda^{NXY}=\lambda^{PXY}=0$	77.47	65
H_4: $H_3 + \lambda^{S}=\lambda^{P}=\lambda^{N}=0$	97.77	68
H_5: $H_3 + \lambda^{N}=\lambda^{S}=0$	79.56	67
H_6: $H_5 + \lambda^{XY}=0$	79.60	68
H_7: $H_6 + \lambda^{XMRY}=\lambda^{XMY}=\lambda^{XRY}=\lambda^{XMR}=0$	88.07	75
H_8: H_7 + Linear Restriction on λ^{XM}	94.44	76

Table 2: Parameter Estimates (and ASE) for Selected Models

Parameter	Model H_0	H_3	H_6	H_7	H_8
λ_1^X	$.430_{(.189)}$	$.546_{(.158)}$	$.575_{(.082)}$	$.542_{(.077)}$	$.528_{(.077)}$
λ_{11}^{XY}	$-.220_{(.190)}$	$.040_{(.049)}$	---	---	---
λ_1^P	$-.221_{(.080)}$	$-.262_{(.075)}$	$-.278_{(.073)}$	$-.278_{(.073)}$	$-.278_{(.072)}$
λ_1^N	$-.026_{(.071)}$	$-.061_{(.061)}$	---	---	---
λ_1^S	$.086_{(.119)}$	$.064_{(.089)}$	---	---	---
λ_{11}^{PY}	$.114_{(.094)}$	---	---	---	---
λ_{11}^{NY}	$.111_{(.072)}$	---	---	---	---
λ_{11}^{SY}	$.084_{(.121)}$	---	---	---	---
λ_{11}^{PX}	$1.417_{(.086)}$	$1.402_{(.077)}$	$1.410_{(.078)}$	$1.409_{(.078)}$	$1.407_{(.078)}$
λ_{11}^{NX}	$1.314_{(.071)}$	$1.294_{(.061)}$	$1.261_{(.054)}$	$1.260_{(.054)}$	$1.258_{(.054)}$
λ_{11}^{SX}	$1.614_{(.117)}$	$1.646_{(.138)}$	$1.674_{(.132)}$	$1.671_{(.131)}$	$1.680_{(.133)}$
λ_{111}^{PXY}	$.060_{(.079)}$	---	---	---	---
λ_{111}^{NXY}	$.051_{(.071)}$	---	---	---	---
λ_{111}^{SXY}	$-.167_{(.119)}$	$-.257_{(.130)}$	$-.256_{(.130)}$	$-.245_{(.130)}$	$-.247_{(.132)}$

Table 2 (cont.): Parameter Estimates (and ASE) for Selected Models

Parameter	Model H_0	H_3	H_6	H_7	H_8
λ_{11}^{XM}	$-.549_{(.082)}$	$-.549_{(.083)}$	$-.557_{(.077)}$	$-.532_{(.047)}$	$-.478_{(.041)}$ [a]
λ_{12}^{XM}	$.088_{(.062)}$	$.083_{(.063)}$	$.086_{(.062)}$	$.106_{(.042)}$	---
λ_{11}^{XR}	$.132_{(.048)}$	$.132_{(.049)}$	$.139_{(.038)}$	$.145_{(.034)}$	$.146_{(.034)}$
λ_{111}^{XRM}	$-.046_{(.082)}$	$-.051_{(.083)}$	$-.061_{(.075)}$	---	---
λ_{121}^{XRM}	$.006_{(.062)}$	$.007_{(.063)}$	$.010_{(.062)}$	---	---
λ_{111}^{XMY}	$-.003_{(.082)}$	$-.002_{(.083)}$	$-.012_{(.074)}$	---	---
λ_{121}^{XMY}	$-.012_{(.062)}$	$-.018_{(.063)}$	$-.014_{(.061)}$	---	---
λ_{111}^{XRY}	$-.011_{(.048)}$	$-.010_{(.049)}$	$.049_{(.039)}$	---	---
λ_{1111}^{XRMY}	$-.024_{(.082)}$	$-.028_{(.083)}$	$-.040_{(.072)}$	---	---
λ_{1121}^{XRMY}	$.093_{(.062)}$	$.093_{(.063)}$	$.098_{(.061)}$	---	---

[a] Estimate for ϕ^{XM}.

Table 3: Estimated Probability of Disapproval of Indicator Variables by Latent Class and Year

Indicator Variable	1971 Class I	1971 Class II	1991 Class I	1991 Class II
Poor Couples	.905	.033	.905	.033
No More Children	.925	.075	.925	.075
Single Women	.946	.054	.979	.021

Table 4: Estimated Probability of Disapproval of Latent Variable (Class I) by Religion and Attitude Toward Premarital Sex

Premarital Sex	Catholic	Protestant
Not Wrong	.597	.452
Sometimes Wrong	.794	.682
Always Wrong	.909	.848

Work and Organization Research Centre
Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, The Netherlands