Growth Mixture Modeling of Adolescent Alcohol Use Data

Terry E. Duncan, Ph.D., Susan C. Duncan, Ph.D.,
Lisa A. Strycker, M.A., Hayrettin Okut, Ph.D., and Fuzhong Li, Ph.D.

Developmental research involves the identification of individual differences in change as well as understanding the process of change itself (Baltes & Nesselroade, 1979; Collins & Horn, 1991). The contemporary approach to the analysis of change, as presented in prior chapters, has focused on growth curve modeling that explicitly considers both intraindividual change and interindividual differences in such change, but treats the data as if collected from a single population. This assumption of homogeneity in the growth parameters is often unrealistic. If heterogeneity is ignored, statistical analyses and their effects can be seriously biased. This paper presents a procedure that accounts for sample heterogeneity – finite mixtures – and their application to longitudinal data.

The underlying theory of finite mixture modeling assumes that the population of interest is not homogeneous (as measured by response probabilities) but consists of heterogeneous subpopulations with varying parameters. Mixture models are well known in the context of latent class models (e.g., Clogg, 1995; Goodman, 1974; Heinen, 1996; Langeheine & Rost, 1988; McCutcheon, 1987).

Clogg (1995) describes latent class models as comprising a subset of the general class of latent structure models including factor analysis models, covariance structure models, latent profile models, and latent trait models. Others approach the mixture problem from a finite mixture distribution perspective (e.g., Everitt & Hand, 1981; Titterington, Smith, & Makov, 1985). The common theme of mixture modeling is to partition the population into an unknown number of latent classes or subpopulations with class membership determined by specific parameters. The purpose of latent class analysis is, therefore, to 1) estimate the number and size of the latent classes in the mixture, 2) estimate the response probabilities for each indicator given the latent class, 3) assign latent class membership to individuals, and 4) examine the fit of the model to the data using various test statistics.

LATENT CLASS ANALYSIS OF DYNAMIC MODELS

When researchers deal explicitly with discrete observed variables, latent class analysis may be used for classifying subjects into a set of mutually exclusive categories (Clogg, 1995; Heinen, 1996). Like other latent variable approaches (e.g., latent variable structural equation modeling), latent class theory is a measurement theory based on the tenant of a static unchanging discontinuous latent variable that divides a population into mutually exclusive and exhaustive latent classes.

On the premise that one needs valid measurement of dynamic latent variables, Collins and colleagues (Collins, 1991; Collins & Cliff, 1990) suggested maintaining a distinction between static and dynamic latent variables, where dynamic latent variables are those that change systematically over time. Within the framework of existing latent class theory, some recent advances have involved formulating latent class models for dynamic processes (e.g., Collins & Wugalter, 1992; Langeheine & Rost, 1988; Meiser & Ohrt, 1996). Among these, the most widely recognized in the area of substance use research is latent transition analysis (LTA: Collins & Wugalter, 1992; Graham, Collins, Wugalter, Chung, & Hansen, 1991). LTA enables the researcher to fit latent class and latent transition models to data by specifying dynamic latent variables that change in systematic ways over time. The analysis can be extended to include multiple indicators for each point in time and has the ability to handle missing data. Collins and colleagues have presented models that allow for tests of stage-sequential development in longitudinal data and have provided applications in the context of alcohol use and other substances (e.g., Collins, Graham, Rousculp, & Hansen, 1997; Collins, Graham, Long, & Hansen, 1994; Graham et al., 1991; Hansen & Graham, 1991).

COVARIANCE STRUCTURE ANALYSIS MIXTURE MODELING

In recent years, mixture models, in the context of covariance structure models (commonly known as structural equation models), have been developed (e.g., Arminger & Stein, 1997; Dolan & van der Mass, 1998; Jedidi, Jagpal, & DeSarbo, 1997; Yung, 1997). Yung (1997) proposed three methods related to finite mixtures of confirmatory factor-analytic (CFA) models with structured means for handling data heterogeneity. Within Yung's mixture approach, observations are assumed to be drawn from mixtures of distinct CFA models. But each observation does not need to be identified as belonging to a particular subgroup prior to model fitting. Yung's approach reduces to regular multigroup CFA under a restrictive sampling scheme in which the structural equation model for each observation is assumed to be known. By assuming a mixture of multivariate normals for the data, Yung utilizes a maximum likelihood estimation approach, the Expectation Maximization (EM) algorithm and the Approximate-Scoring method, to fit models.

Arminger and Stein (1997) presented a mixture model that includes CFA models and structural regression models among latent variables. Their model allows for the introduction of fixed observed regressors (predictors), such as gender. The inclusion of fixed observed regressors allows one to replace the requirement of unconditional normality with the requirement of conditional normality within each component of the mixture. This approach allows the researcher to specify both conditional and unconditional normal mixtures subject to structural equation modeling. The expected values and covariance matrices of the mixture components are parameterized using conditional mean and covariance structures. The likelihood function consisting of the EM algorithm and a weighted least squares loss function is constructed to test the fit of the models.

Extending others' work in this area (Jedidi et al., 1997; Yung, 1997), Dolan and van der Mass (1998) recently proposed an approach to covariance and mean structure modeling within unconditional multivariate normal mixtures that includes common CFA models and structural equation models. Like other approaches, Dolan et al.'s method is applicable in situations where the component membership of the sample is unknown, but allows linear and nonlinear constraints to be imposed on the parameters in the model. That is, within the general model, individual parameters can be subjected to equality, nonlinear, and simple bounds constraints. These constraints enable one to investigate a variety of substantive hypotheses concerning the differences between the components in the mixture.

GROWTH MIXTURE MODELING

Recently, researchers have begun exploring new ways of constructing more complex and dynamic models that are better suited for assessing change (Collins & Sayer, 2001). Most recently, Muthén (2001) proposed an extension of current LGM methodology that includes relatively unexplored mixture models, such as growth mixture models, mixture structural equation models, and models that combine latent class analysis and structural equation modeling.

Relevant to longitudinal research is the growth mixture modeling approach, combining categorical and continuous latent variables into the same model. Muthén and colleagues (Muthén, Brown, Khoo, Yang, & Jo, 1998; Muthén & Muthén, 1998; Muthén & Shedden, 1999) described in detail the generalization of LGM to finite-mixture latent trajectory models and proposed a general growth mixture modeling framework (GGMM). The GGMM approach allows for unobserved heterogeneity in the sample, where different individuals can belong to different subpopulations. The modeling approach provides for the joint estimation of (a) a conventional finite mixture growth model where different growth trajectories can be captured by class-varying random coefficient means and (b) a logistic regression of outcome variables on the class trajectory. The model can be further extended to estimate varying class membership probability as a function of a set of covariates (i.e., for each class the values of the latent growth parameters are allowed to be influenced by covariates) and to incorporate outcomes of the latent class variable.

Figure 1 displays a full growth mixture model within the framework of Muthén and colleagues. This model contains a combination of a continuous latent growth variable, η_j (j = Intercept and Slope) and a latent categorical variable, C, with K classes, $C_i = (c_1, c_2, \ldots, c_K)'$, where $c_i = 1$ if individual i belongs to class k and zero otherwise. These latent attributes are represented by circles in Figure 1. The latent continuous growth variable portion of the model represents conventional growth modeling with multiple indicators, Y, measured at four time points (e.g., Willett & Sayer, 1994). The categorical latent variable is used to represent latent trajectory classes underlying the latent growth variable, η. Both latent continuous and latent class variables can be predicted from a set of background variables or covariates, X, since the model allows the mixing proportions to depend on prior information and/or subject-specific variables. The growth mixture portion of the model, on the other hand, can have mixture outcome indicators, U. In this model, the directional arrow from the latent trajectory classes to the growth factors indicates that the intercepts of the regressions of the growth factors on X vary across the classes of C. The directional arrow from C to U indicates that the probabilities of U vary across the classes of *C*.

Muthén (2001) considers growth mixture modeling as a second generation of structural equation modeling. Indeed, the general framework outlined by Muthén (1998) provides new opportunities for growth modeling. Growth mixture models are applicable to longitudinal studies where individual growth trajectories are heterogeneous and belong to a finite number of unobserved groups. The application of mixtures to growth modeling also may be used as an alternative to cluster analytic techniques if the posterior probability of membership of an individual in a latent class is used to assign latent class membership.

Examples of further extensions of this application may include

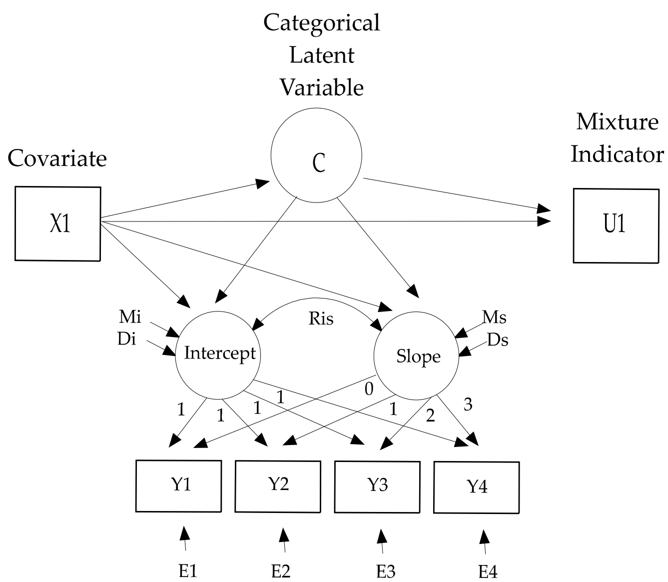

FIG. 1. Representation of the growth mixture model.

mixtures of multivariate growth models, where several simultaneous growth processes are present, and multigroup growth mixture models where different latent growth structures exist in different subgroups that are assumed to be independently sampled. In the former, the researcher specifies multiple-outcome growth mixture models to study multivariate change, which has been described by Duncan and Duncan (1996), Duncan et al., (1999), and Tisak and Meredith (1990). In the latter, the researcher estimates models across known groups using cross-group constraints and makes comparisons based on the parameter values and latent class predictions (e.g., Clogg, 1995). Considerable methodological and applied work needs to be done to fully explore the utility of growth mixture modeling.

MODEL SPECIFICATIONS

Let π_{ik}=P(C_i=k) denote the probability that the i^{th} response falls in the k^{th} category. For example π_{i1} is the probability that the i^{th} response belongs to the category 1 ($\sum_{i=1}^{K} \pi_{i1} = 1$ for i, where the probabilities add up to one for each individual and there are K-1 parameters).

In a *multinomial* distribution, the probability distribution of the counts c_{ij} given the total N is

$$p\left\{C_{i1} = c_{i1}, C_{i2} = c_{i2} \ldots C_{iK} = c_{iK}\right\} = \binom{N}{c_{i1}, \ldots, c_{iJ}} \pi_{i1}^{c_{i1}} \ldots \pi_{iK}^{c_{iK}} . \qquad (1)$$

Equation 1 becomes *binomial* in the case of K=2. In terms of response probabilities, the model for the *multinomial logit* is

$$\pi(c_i) = \frac{\exp\{\Delta_{ik}\}}{\sum_{K-1}^{K} \exp\{\Delta_{ik}\}} \qquad (2)$$

where $\Delta_{i1} = \log \frac{\pi_{i1}}{\pi_{iK}}$.

Let u represent a binary outcome of the latent class analysis model (LCA) and c represent the categorical latent variable with K classes. The individual (marginal) probability density of U_I *is*

$$P(u_i = 1) = \sum_{K-1}^{K} P(c = k) P(u_i = 1 | c = k) . \qquad (3)$$

If the model relates c to covariate x by means of a logit multinomial model, Equation 3 would be:

$$P(c_{ik} = 1|x_i) = \frac{\exp(a_{ck} + \gamma_{ck} x_i)}{\sum_{K-1}^{K} \exp(a_{ck} + \gamma_{ck} x_i)} \quad \text{where} \begin{cases} c_{ik} = 1 \ \textbf{if} \ c_{ik} \in K \\ \textbf{otherwise} \ c_{ik} = 0 \end{cases} . \tag{4}$$

A multinomial logistic regression uses the standardization of zero coefficients for $a_{ck} = 0$ and $\gamma_k = 0$. Therefore, the logit c for the odds of class K is

$logit\ c = log[P(c_i{=}k \mid x_i)/P(c_i{=}K \mid x_i)] = a_{ck} + \gamma_k x_i$.

Consider the model presented in Figure 1 in which the outcome variable is measured repeatedly across four time points (y_1, y_2, y_3, y_4). In addition, two latent growth factors and three latent classes (C = 1,2,3) are postulated, with one covariate (*X1*) and one mixture indicator (*U1*).

The following equation defines the conventional latent growth model for the continuous observed variables Y with continuous latent variables η for individual i across 3 years of data:

$$Y = \Lambda\eta + \Theta \tag{5a}$$

where Y is a vector ($Y' = y_{i1}, y_{i2}, ..., y_{it}$) containing scores for individual i (i = 1, 2, ..., N) at t (t = 1, 2, ..., m) occasions, η is defined as a $px1$ vector of intercept and linear factors, and Λ is an mxp design matrix (or basis functions; Meredith & Tisak, 1990) representing specific aspects of change. Based on information from Figure 1, m=4, p=2, and Λ and η are 4x2 and 2x1 matrices, respectively. Column 1 of Λ is defined as the intercept factor by fixing all loadings at 1.0. Column 2 is defined as the slope (linear rate of change) by setting the loadings λ_{12}, ..., λ_{42} equal to the values of yearly measurement (t = 0, 1, 2, 3) for individual i (Note that λ_{12} is set to 0 so that the intercept can be interpreted as the predicted value of the response variable at the first year of measurement). Θ is a 4 x 1 vector of residual terms for individual i. The model expressed in Equation 5a has the matrix form:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1 & \lambda_{12} \\ 1 & \lambda_{22} \\ 1 & \lambda_{32} \\ 1 & \lambda_{42} \end{bmatrix} \begin{bmatrix} \eta_0 \\ \eta_s \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{bmatrix} \tag{5b}$$

In fitting this model, estimates of the factor loadings in Λ, the variances and covariances of the latent factors (η_0, η_s) in Φ, and the means (α_0, α_s) of the latent factor scores in η can be obtained.

The two different curve shapes are hypothesized to be captured by class-varying random coefficient means for each *K-1* class, $\alpha_c[\eta_0, \eta_s]$. When including the effect of covariates in the model the continuous latent variables η for individual *i* are linked to the categorical latent variable *C* and to the observed covariant vector *X* (where $X = x_1, x_2, ..., x_N$) through the following equation:

$$\eta_i = A_{Ci} + \Gamma_\eta X_i + \zeta_i \tag{6}$$

where A_{Ci} equals an mxK matrix containing columns α_k, $k = 1, 2, ..., K$, of intercept parameters of each *C* class, Γ_η is an mxp parameter matrix containing the effects of *X* on η given *C* categories, and ζ is an *m*-dimensional residual vector, normally distributed, uncorrelated with other variables, and with zero mean and covariance matrix Ψ^k of

$$\Psi^k = \begin{bmatrix} \Psi_0^k & \\ \Psi_{0s}^k & \Psi_s^k \end{bmatrix}$$

Binary outcome variables, *U*, contain both a measurement model and a structural model (Muthén & Muthén, 1998). In the measurement portion of the model, the r binary variables U_{ij} are assumed to be conditionally independent given C_i and X_i, with the following conditional probability decomposition:

$$P(U_{Ij} \mid C_i, X_i) = P(u_{i1}, u_{i2}, ..., u_{ir} \mid C_i, X_i) = P(u_{i1} \mid C_i, X_i)\, P(u_{i2} \mid C_i, X_i) ... P(u_{ir} \mid C_i, X_i) \tag{7}$$

We define $\tau_{ij} = P(u_{ij} = 1 \mid C_i, X_i)$, the r-dimensional vector $\tau_i = (\tau_{i1}, \tau_{i2}, ..., \tau_{ir})'$, and the r-dimensional vector logit $(\tau_i) = (\log[\tau_{i1} / (1 - \tau_{i1})], \log[\tau_{i2} / (1 - \tau_{i2})],$

..., $\log[\tau_{ir} / (1 - \tau_{ir})])'$. The logit model is, therefore:

$$\text{logit } (\tau_i) = \Lambda_u C_i + K_u X_i, \tag{8}$$

where Λ_u is an rxK parameter matrix and K_u is an rxq parameter matrix and $\tau_{ijk} = P(U_{ijk} = 1 \mid C_{ik} = 1)$. For the model in Figure 1, where $U = (U1)$, $X = (X1)$, and $C = 1,2,3$, the matrices of Λ_u and K_u would be

$$\Lambda_u = \left[\lambda_{11,12,13}\right]', \quad K_u = \left[k_{11}\right].$$

In the structural portion of the model, the categorical latent variables of C represent mixture components that are related to X through a multinomial logit regression model for an unordered polytomous response. Defining $\pi_{ik} = P(C_{ik} = 1 \mid X_i)$, the K-dimensional vector $\pi_i = (\pi_{i1}, \pi_{i2}, ..., \pi_{iK})'$, and the K - 1-dimensional vector logit $(\pi_i) = (\log[\pi_{i1} / \pi_{iK}], \log[\pi_{i2} / \pi_{iK}], ..., \log[\pi_{i, K-1} / \pi_{iK}])'$,

$$\text{logit } (\pi_i) = \alpha_C + \Gamma_C X_i, \tag{9}$$

where α_C is a K-1dimensional parameter vector and Γ_C is a $(K - 1) \times p$ parameter matrix. Again, letting $C = 1,2,3$ and $X = (X1)$, α and Γ of the latent class regression model part of the model are

$$\alpha = \left[\alpha_{c1,c2}\right]', \quad \Gamma = \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \end{bmatrix},$$

where π is a 1x1 vector containing the regression intercept and Γ is a 1x1 matrix containing the regression coefficient.

In the model of Equations 5 through 9, the finite mixture arises because the conditional distribution of Y and U given X is governed by parameters that vary across the categories of C; the mean vector of Y is allowed to vary because of the inclusion of C in Equation 7 and the probabilities of U are allowed to vary because of the inclusion of C in Equation 9.

MODEL IDENTIFICATION AND ESTIMATION

Identification of parameters in finite mixture modeling is discussed in Muthén and Shedden (1999). The authors point out that the model is not identified without some parameter restrictions which have to be determined for a given application. The covariance matrices Ψ and Θ are class-invariant while the random coefficient η means of A vary across classes with class probabilities determined by α_c. Identification, therefore, concerns a multivariate normal mixture for y with a class-invariant covariance matrix and with means that are functions of the class-varying means of the reduced dimension m of the underlying η. In addition, Muthén and Muthén (1998) suggested two guides for signs of empirical identification: (a) the invertibility of the information matrix used to produce the standard errors and (b) observing a change in the log likelihood when freeing a parameter in an identified model.

The primary method used to estimate parameters in finite mixture modeling is a maximum likelihood (ML) technique, specifically, the Expectation Maximization (EM) algorithm. In the EM algorithm, estimation involves the calculation of an E-step and an M-step. The E-step evaluates conditional expectations of various complete-data sufficient statistics. The M-step updates unknown parameters by maximizing the conditional expectations obtained in the E-step. EM is efficient even when estimates are far from optimum. Other algorithms such as Fisher-Scoring and Newton-Raphson are more efficient for rapid convergence. In some applications, the ideal solution to estimation is to combine the advantages of two different methods. For example, parameter estimation would start with a number of EM iterations and then switch to Fisher-Scoring or Newton-Raphson (using information or Hessian matrix) when estimates are close to those of the final solution. Mplus (Muthén & Muthén, 1998) uses the principle of ML estimation and employs the algorithm of EM for maximization and bootstrapping standard errors. A thorough description of the procedure is given in Muthén and Shedden (1999).

Model evaluation for growth mixture models proceeds much like in conventional SEM or latent growth models for homogeneous populations. Model fit for a mixture analysis is performed by the log likelihood value. Using chi-square-based statistics (i.e., the log likelihood ratio), fit for nested models can be examined. It is, however, not appropriate to use such values for comparing models with different numbers of classes, given that this involves inadmissible parameter values of zero class probabilities. In these instances, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) can be used instead. Mplus also provides a sample-size adjusted BIC (ABIC) which has been shown to give superior performance in a simulation study for latent class analysis models (Yang, 1998).

EXAMPLE 1: THE SINGLE-CLASS GROWTH CURVE MODEL

In modeling growth mixture models, a step-by-step approach is recommended: (a) unconditional analyses, followed by (b) conditional analyses. The unconditional analyses specify the model to include no predictor variables or mixture indicators for each class, but with class-invariant and/or class-varying mean and covariance structure. The conditional analysis takes into account influences from predictor variables and subsequent mixture indicators. An underlying assumption of the conventional growth model is that the data come from a single-population growth model that encompasses different types of trajectories (i.e., all individuals belong to one and the same population). It is also assumed that the covariates (X) have the same influence on the growth factors for all trajectories. Data from 206 participants of the National Youth Survey (NYS) with complete data were used in the following examples. Figure 2 represents the single-class mixture model. Input statements for the single class mixture model are presented in Input 1.

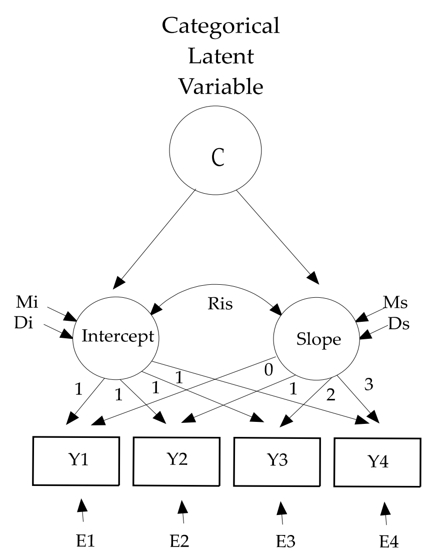

FIG 2. The single-class mixture model.

INPUT 1
MPlus Input Statements for the Single Latent Class Mixture Model

```
Title: Mixture modeling: 1 latent class [no heterogeniety]
Data: File is mixexample.dat;
Variable: Names are id x1 y1 y2 y3 y4 u1;
Usevariables are y1 y2 y3 y4;
Classes = c(1);
Analysis: Type = mixture;
          miteration = 300;
Model:    %overall%
          int by y1-y4@1;
          slp by y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0];
          y1 y2 y3 y4;
          int*.109 slp*;
          int with slp*;
          ! latent class designation %c#1%
          %c#1%
          [int*4.096 slp*.843];
          y1 y2 y3 y4;
          int*1.986 slp*.160;
          int with slp*-.531;
Output: tech1 tech7;
```

Model fitting procedures for the single-class mixture model resulted in a log likelihood H0 value of -1547.067 and a BIC of 3142.085. Significant intercept Mi = 2.850, t = 21.823, p < .001, and slope Ms = .558, t = 12.143, p < .001 means indicated that significant levels of alcohol use were observed at T1, and significant growth in alcohol use occurred over the four assessments. The intercept variance for adolescent alcohol use, Di = 2.607, t = 7.737, p < .01, and the variance of the latent slope scores, Ds = .174, t = 3.504, p < .01, indicate substantial variation existed among individuals in initial status and growth of alcohol use.

Reproduced means are shown in Figure 3. Examination of the reproduced means suggested considerable linearity in alcohol use over time for the single-class model.

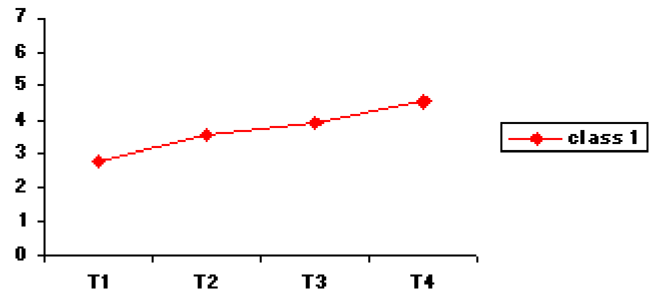

FIG. 3. Single-class model mean trend.

EXAMPLE 2: DETERMINING SAMPLE HETEROGENEITY: MULTIPLE-CLASS GROWTH CURVE MODELS

Specifications for the two-class mixture model are presented in Input 2. Key elements in the model specifications are the statement *Classes = c(2)*; the specification of the second latent class, *%c#2%*; and the varying parameter estimates for the mean intercept and slope, int*2.758 slp*.516, the variances for the intercept and slope, int*1.418 slp*.240, and covariance between the two growth parameters, int with slp*-.593.

Model fitting procedures for the two-class mixture model resulted in a log likelihood H0 value of -1496.325, a BIC of 3093.879, and an entropy estimate of .796. For class 1, estimates for the intercept Mi = 4.442, t = 12.458, $p < .001$, and slope Ms = .720, t = 5.968, $p < .001$ means were significant. Estimates for the latent intercept, Ds = 1.278, t = 1.679, $p < .01$, and slope variance, Di = .174, t = 2.506, $p < .01$, also were significant. The covariance between the intercept and slope factors was not significant, R = -.390, t = -1.686, $p > .05$. For class 2, estimates for the intercept Mi = 1.881, t = 14.071, $p < .001$, and slope Ms = .484, t = 5.002, $p < .01$ means were significant. Estimates for the latent intercept, Ds = 1.081, t = 2.194, $p < .01$, and slope variance, Di = .232, t = 3.493, $p < .01$, also were significant. The covariance between the intercept and slope factors was not significant, R = -.134, t = -.814, $p > .05$.

INPUT 2
MPlus Input Statements for the Two Latent Class Mixture Model

```
Title: Mixture modeling: 2 latent classes
Data: File is mixexample.dat;
Variable: Names are id x1 y1 y2 y3 y4 u1;
Usevariables are y1 y2 y3 y4;
Classes = c(2);
Analysis: Type = mixture;
          miteration = 300;
Model:    %overall%
           int by y1-y4@1;
          slp by y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0];
          y1 y2 y3 y4;
          int*.109 slp*;
          int with slp*;
          [c#1*0];
          ! latent class designation %c#1%
          %c#1%
          [int*4.096 slp*.843];
          y1 y2 y3 y4;
          int*1.986 slp*.160;
          int with slp*-.531;
          ! latent class designation %c#2%
          %c#2%
          [int*2.758 slp*.516];
          y1 y2 y3 y4;
          int*1.418 slp*.240;
          int with slp*-.593;
Output: tech1 tech7;
```

Reproduced means for the two-class model are shown in Figure 4. The mean trends shown in Figure 4 graphically depict the greater mean levels in intercept and slope scores for latent class 1 compared to latent class 2. Final class counts and proportion of total sample size for the two latent class model were 76.07446(0.369) for class 1 and 129.92554(.631) for class 2. Classification of individuals based on their most likely class membership resulted in class counts and proportions of 78(.379) for class 1 and 128(.621) for class 2.

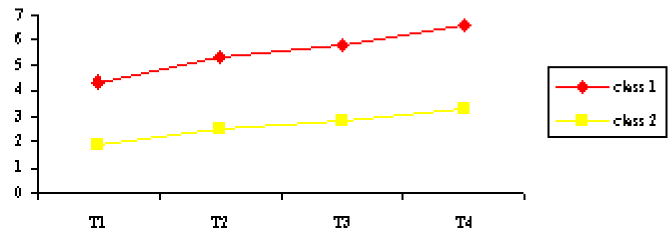

FIG. 4. 2-class model mean trend

Given the significant means and variances derived from the two-class model, the viability of a three-class solution was examined. Input specifications for the three-class mixture model are presented in Input 3. Note that the inclusion of the third class is accomplished by including the statement,*%c#3%*, in the parameter specifications.

Model fitting procedures for the three-class mixture model resulted in a log likelihood H0 value of -1361.258, a BIC of 2877.024, and an entropy estimate of .847. For class 1, estimates for the intercept, $Mi = 4.096$, $t = 8.725$, $p < .001$, and slope, $Ms = .843$, $t = 9.947$, $p < .001$, means were significant. Estimates for the latent intercept, $Ds = 1.986$, $t = 3.182$, $p < .01$, and slope, $Di = .160$, $t = 2.710$, $p < .01$, variance also were significant. The covariance between the intercept and slope factors was significant, $R = -.531$, $t = -3.746$, $p < .01$.

For class 2, estimates for the intercept $Mi = 2.758$, $t = 5.536$, $p < .001$, and slope $Ms = .516$, $t = 1.973$, $p < .01$ means were significant. Estimates for the latent intercept, $Ds = 1.418$, $t = 1.692$, $p < .05$, and slope, $Di = .240$, $t = 1.663$, $p < .01$, variance also were significant. The covariance between the intercept and slope factors was not significant, $R = -.593$, $t = -1.751$, $p > .05$.

For class 3, the estimate for the intercept mean $Mi = 1.233$, $t = 19.189$, $p < .001$, was significant while the estimate for the slope mean $Ms = .180$, $t = 1.174$, $p > .05$, was not significant. The estimate for the latent intercept variance, $Ds = .109$, $t = 2.068$, $p < .05$, was significant whereas the slope variance, Di = .053, t = .577, p > .05, was not significant. The covariance between the intercept and slope factors was not significant, R=-.014,t= -.497, p > .05. Class counts and proportion of total sample size for the three latent class model were 83 (40%), 62 (30%)and 61 (30%) for class 1, 2, and 3, respectively (Appendix 2).

INPUT 3

MPlus Input Statements for the Three Latent Class Mixture Model

```
Title: Mixture modeling: 3 latent classes
Data: File is mixexample.dat;
Variable: Names are id x1 y1 y2 y3 y4 u1;
Usevariables are y1 y2 y3 y4;
Classes = c(3);
Analysis: Type = mixture;
          miteration = 300;
Model:    %overall%
          int by y1-y4@1;
          slp by y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0];
          y1 y2 y3 y4;
          int*.109 slp*;
          int with slp*;
          [c#1*0];
          ! latent class designation %c#1%
          %c#1%
          [int*4.096 slp*.843];
          y1 y2 y3 y4;
          int*1.986 slp*.160;
          int with slp*-.531;
          %c#2%
          [int*2.758 slp*.516];
          y1 y2 y3 y4;
          int*1.418 slp*.240;
          int with slp*-.593;
          %c#3%
          [int*1.233 slp*];
          int*1.018 slp*.140;
          int with slp*-.193;
Output: tech1 tech7;
```

Reproduced means for the three-class model are shown in Figure 5.

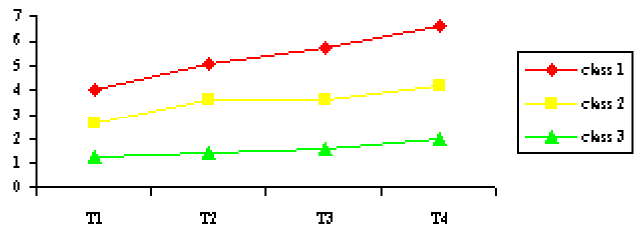

FIG. 5.-class model mean trend.

CHOOSING AN APPROPRIATE NUMBER OF LATENT CLASSES

A test of model fit can be obtained from the log likelihood value for a given model. For comparing models with different numbers of classes the log likelihood value is not appropriate, but the AIC and BIC information criteria can be used instead. The sample size adjusted BIC was found to be superior in a simulation study for latent class analysis (Yang, 1998). The sample size adjusted BIC values for the competing models are presented in Figure 6.

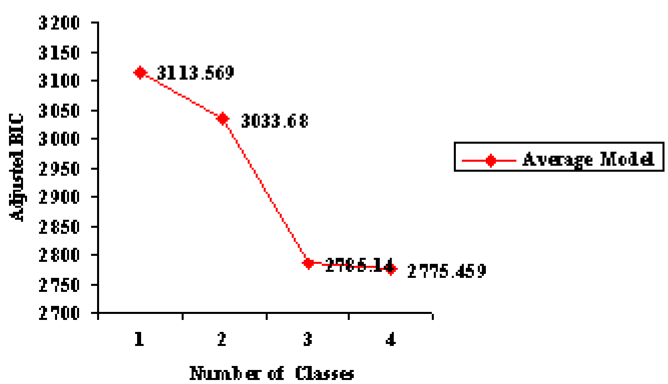

FIG. 6. BIC function.

The degree to which the latent classes are clearly distinguishable can be assessed by using the estimated posterior probabilities for each individual. By classifying each individual into their most likely class, a table can be created with rows corresponding to individuals classified into a given class and column entries representing the average conditional probabilities (Nagin, 1999). Diagonals close to 1 and off-diagonals close to 0 represent good classification rates. A summary measure of the classification is given by the entropy measure (Ramaswamy et al., 1993) where

$$E_k = 1 - \frac{\sum_i \sum_k (-\hat{p}_{ik} \ln \hat{p}_{ik})}{n \ln K} \tag{10}$$

and $\hat{p}_{ik}$ denotes the estimated conditional probability for individual i in class k. Entropy values range from 0 to 1, with values close to 1 indicating greater clarity in classification. Entropy values for the competing models are presented in Figure 7.

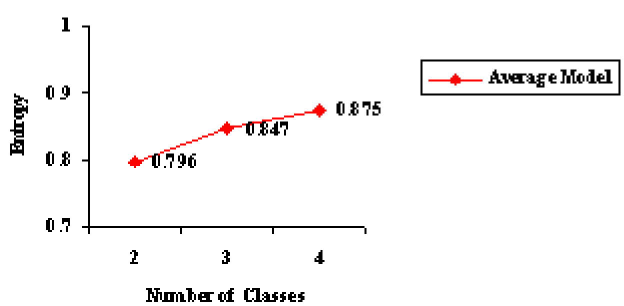

FIG. 7. Entropy values.

EXAMPLE 3: INCLUDING COVARIATES

Having described the basic unconditional mixture model in the prior sections, we now examine the effects of the covariate X1 on the latent growth parameters using the three-class model previously tested. Figure 8 presents the conditional growth mixture model which includes the effects of the covariate, X1, on the growth parameters.

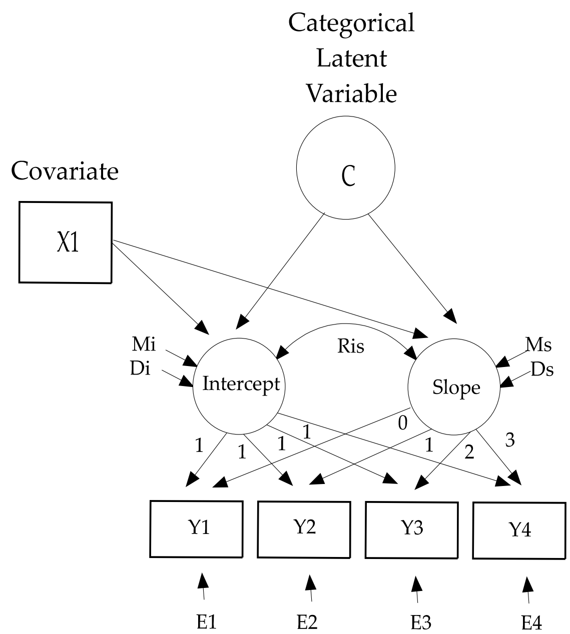

FIG. 8. Effects of the covariate on the growth parameters.

Input specifications for the effects of the covariate on the growth factors are presented in Input 4. Key specifications of the regression of the growth parameters on X are given by the input statements *int on x1*; and *slp on x1*;.

Model fitting procedures for the three-class conditional mixture model resulted in a log likelihood H0 value of -1339.043, a BIC of 2864.561, and an entropy estimate of .829. For class 1, the estimate for the intercept mean, Mi = -2.468, t = -1.745, $p > .05$, was not significant, whereas the estimate for the slope mean, Ms = 2.526, t = 4. 99, $p < .01$ was significant. Estimates for the latent intercept, Ds = 1.131, t = 2.013, $p < .05$, and slope, Di = .130, t = 2.072, $p < .05$, variance also were significant. The covariance between the intercept and slope factors was significant, R = -0.345, t = -2.004, $p < .05$.

INPUT 4

MPlus Input Statements for the Effects of the Covariate on the Growth Factors

```
Title: Mixture modeling: 3 latent classes F on X
Data: File is mixexample.dat;
Variable: Names are id x1 y1 y2 y3 y4 u1;
Usevariables are x1 y1 y2 y3 y4;
Classes = c(3);
Analysis: Type = mixture;
          miteration = 300;
Model:    %overall%
          int by y1-y4@1;
          slp by y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0];
          y1 y2 y3 y4;
          int*.109 slp*0;
          int with slp*0;
          ! effect of covariate on growth factors     [constrained in the overall model]
           int on x1;
          slp on x1;
          [c#1*0];
          %c#1%
          [int*4.096 slp*.843];
          y1 y2 y3 y4;
          int*1.986 slp*.160;
          int with slp*-.531;
          ! effect of covariate on growth factors     [freely estimated in  each class]
          int on x1;
          slp on x1;
          %c#2%
          [int*2.758 slp*.516];
          y1 y2 y3 y4;
          int*1.418 slp*.240;
          int with slp*-.593;
          ! effect of covariate on growth factors     [freely estimated in each class]
           int on x1;
          slp on x1;
          %c#3%
          [int*1.233 slp*0];
          int*1.018 slp*.140;
          int with slp*-.193;
          ! effect of covariate on growth factors     [freely estimated in  each class]
          int on x1;
          slp on x1;
Output: tech1 tech7;
```

For class 2, estimates for the intercept, *Mi* = .479, t = .358, p > .05, and slope, *Ms* = .142, t = .264, p > .05, means were not significant, indicating that no significant alcohol use was observed at T1 and no significant growth in alcohol use occurred over the four assessments for these individuals. Estimates for the latent intercept, *Ds* = 1.130, t = 1.367, p > .05, and slope, *Di* = .214, t = 1.607, p > .05, variance also were not significant. The covariance between the intercept and slope factors was not significant, R = -.497, t = -1.574, p > .05. For class 3, the estimates for the intercept, *Mi* = .451, t = .771, p > .05, and slope, *Ms* = -.083, t = -.04, p > .05, means were not significant. The estimate for the latent intercept, *Ds* = .116, t = 1.076, p > .05, and slope, *Di* = .068, t = .364, p > .05, variance also were not significant. The covariance between the intercept and slope factors was not significant, R = -.015, t = -.444, p > .05.

For class 1, the coefficients for the regression of the growth factors on *X1* (Figure 7) were 1.794, t = 4.965, p < .01, for the intercept, and -.459, t = -3.424 and p < .01 for the slope, indicating a significant effect of general deviant behavior (X1 = minor assaults, public disorder, property damage, and minor theft) on the intercept and slope of alcohol use where greater levels of general deviancy were associated with elevated levels of alcohol use at T1 and a significant reduction in alcohol use trajectories over the four assessments.

The regression coefficients for class 2 and class 3 were all nonsignificant. For class 2, the regression coefficients were .659, t = 1.486, p > .05 and .124, t = .724, p > .05 for the intercept and slope, respectively. For class 3, the regression coefficients were .266, t = 1.218, p > .05 for the intercept, and .089, t = .113, p > .05 for the slope of alcohol use. The regression coefficients on *X1* were nonsignificant for both latent classes (class 2 and class 3).

The BIC (2.864.56) improved as a result of including the covariate in the conditional model. In addition, with the inclusion of the covariate in the conditional model, the proportions for each class changed from .40, .30, and .30 to .38, .32, and .30 for classes 1, 2, and 3, respectively.

Figure 9 presents the conditional growth mixture model which includes the effects of the covariate, X1, on the growth parameters and on the latent class variable, C.

Input specifications for the effects of the covariate on the latent classes, C, are presented in Input 5. Because C is an unordered categorical latent

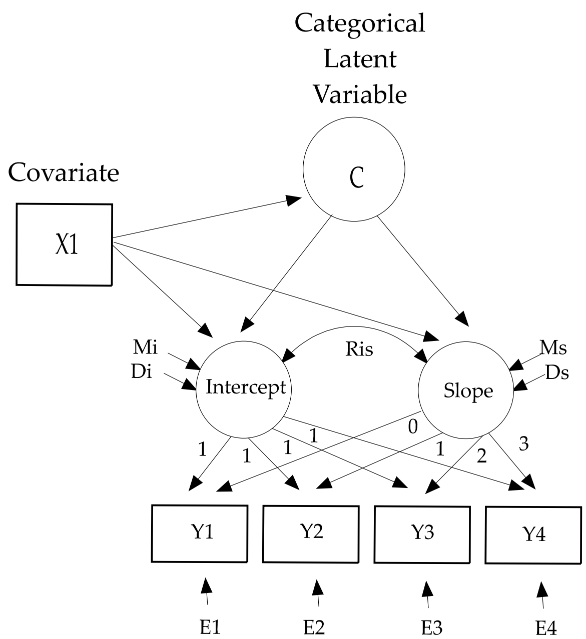

FIG. 9. Effects of the covariate on the growth parameters and latent classes(C).

variable with three categories, the interpretation of the effect of the regression of C on X is not the same as for other latent variable models using continuous latent variables. Instead, it represents the multinomial logistic regression of C on X1. Regression equations are specified for the first two classes while the intercept (latent class mean) and slope (logistic regression coefficient) for the last class are fixed at zero as the default. Key specifications of the regression of the latent classes on X are given by the input statements *c#1 on x1*; and *c#2 on x1*; where the number of latent classes specified are defined by C-1 classes. Note that these specifications, unlike those for the regression of the growth parameters on X, are specified only in the overall model.

INPUT 5

MPlus Input Statements for the Effects of the Covariate on the Latent Classes

```
Title: Mixture modeling: 3 latent classes C on X
Data: File is mixexample.dat;
Variable: Names are id x1 y1 y2 y3 y4 u1;
Usevariables are x1 y1 y2 y3 y4;
Classes = c(3);
Analysis: Type = mixture;
          miteration = 300;
Model:    %overall%
          int by y1-y4@1;
          slp by y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0];
          y1 y2 y3 y4;
          int*.109 slp*0;
          int with slp*0;
          int on x1;
          slp on x1;
          [c#1*0];
          !c regressed on x [overall model only]
          c#1 on x1;
          c#2 on x1;
          %c#1%
          [int*4.096 slp*.843];
          y1 y2 y3 y4;
          int*1.986 slp*.160;
          int with slp*-.531;
          int on x1;
          slp on x1;
          %c#2%
          [int*2.758 slp*.516];
          y1 y2 y3 y4;
          int*1.418 slp*.240;
          int with slp*-.593;
           int on x1;
          slp on x1;
          %c#3%
          [int*1.233 slp*0];
          int*1.018 slp*.140;
          int with slp*-.193;
          int on x1;
          slp on x1;
 Output: tech1 tech7;
```

Model fitting procedures resulted in a log likelihood H0 value of -1306.417 and entropy estimate of .860. In addition, the BIC value improved (2809.965) as a result of including the regression of both latent classes and growth factors (intercept and slope) on X1.

In this model, latent class membership was regressed on the covariate (X1). The logistic regression coefficients for the effects of X1 on class 1 and 2 were 3.185, $t = 6.115$, $p < .01$, and 2.158, $t = 4.020$, $p < .01$, respectively. The logistic regression coefficient for the last class was fixed at zero to serve as the reference category.

These positive logistic regression coefficients (3.185 and 2.158) indicate that the logit of the probability for class 1 and class 2 increases as deviancy in behavior increases. In multinomial logits, one of the classes serves as the baseline or reference category, in this case, class 3. Therefore, we calculate logits for the other categories (class 1 and class 2) relative to the reference category and then let the logits be a linear function of the predictors. The missing contrast between 1 and 2 can easily be obtained in terms of the other two, since

$$\Delta_{ik} = \log\frac{\pi_{i1}}{\pi_{i2}} = \log\frac{\pi_{i1}}{\pi_{i3}} - \log\frac{\pi_{i2}}{\pi_{i3}} \quad . \tag{11}$$

Here we need only K-1 equations (see Equation 2 and Equation 3) to describe a variable with K response categories.

Intercepts for the latent classes were estimated at values of -10.307, $t = -6.109$, $p < .01$ and -6.778, $t = -3.888$, $p < .01$, for classes 1 and 2, respectively. The intercept for the last class was fixed at zero to serve as the reference category.

For class 1, the coefficients of the regression on *X1* (Figure 9) for the intercept, 1.576, $t = 4.885$, $p < .01$, and slope, -.420 $t = -3.544$, $p < .01$, were significant. The coefficient of regression on *X1* for intercept and slope for class 2, .359, $t = .934$, $p > .05$ and .082, $t = .430$, $p > .05$, were both nonsignificant. For class 3, estimates, .262, $t = 1.845$, $p > .05$ and -.137, $t = -1.376$, $p > .05$, also were nonsignificant.

For class 1, the estimate for the intercept mean, $Mi = -1.778$, $t = -1.458$, $p > .05$, was not significant whereas the estimate of the slope mean, $Ms = 2.421$, $t = 5.339$, $p < .001$, was significant. Estimates for the latent intercept, $Di = 1.290$, $t = 3.233$, $p < .01$, and slope, $Ds = .131$, $t = 2.405$, $p < .05$, variance also were significant. The covariance between the intercept and slope factors was significant, $R = -0.381$, $t = -2.824$, $p < .01$.

For class 2, estimates for the intercept, $Mi = 1.553$, $t = 1.093$, $p > .05$, and slope, $Ms = 0.254$, $t = 0.384$, $p > .05$, means were not significant. The estimate for the latent intercept variance, $Di = 1.280$, $t = 1.512$, $p > .05$, was not significant but the estimate for the slope variance, $Ds = .245$, $t = 1.1732$, $p < .05$, was significant. The covariance between the intercept and slope factors was significant, $R = -.582$, $t = -1.712$, $p < .05$.

For class 3, the estimates for the intercept, $Mi = .465$, $t = 1.169$, $p > .05$, and slope, $Ms = .602$, $t = 1.841$, $p > .05$, means were not significant. The estimates for the latent intercept, $Di = .089$, $t = 2.186$, $p < .05$, and the slope, $Ds = .058$, $t = 1.969$, $p < .05$, variance were significant. The covariance between the intercept and slope factors was not significant, $R = -.004$, $t = -.163$, $p > .05$.

Classification rates changed once again with the inclusion of the effects of the covariate on the latent class variable. When the model includes a covariate, the probability that individual i falls in class k of the latent class variable c is expressed as the multinomial logistic regression given in Equation 4. The proportion of individuals in classes 1, 2, and 3 were .41, .29, and .30, respectively.

EXAMPLE 4: INCLUDING MIXTURE INDICATORS.

Figure 10 presents the conditional growth mixture model which includes the effects of the categorical latent class variable, C, on the mixture indicator, U.

In this model, U represents a distal outcome predicted by class membership. The effect of U on C indicates that the probabilities of U vary across the classes of C. Input specifications for the effects of the categorical latent class variable on the mixture indicator are presented in Input 6. Key specifications of the logistic regression of the mixture indicator on the latentclasses are given by the input statement *[u1$1];* which can include starting values in the logit scale for the threshold of the latent class indicator. The thresholds define the conditional probabilities of U for each class. Starting values are required for the estimation of these thresholds. In our example, the starting values for U in classes 1, 2, and 3, are 1, .5, and 0, respectively. Note that the specification, *categorical=u1*; is included to define U as a categorical mixture indicator. The regressions of U on C are specified as varying parameters in each of the latent classes and appear in each of the latent class specifications.

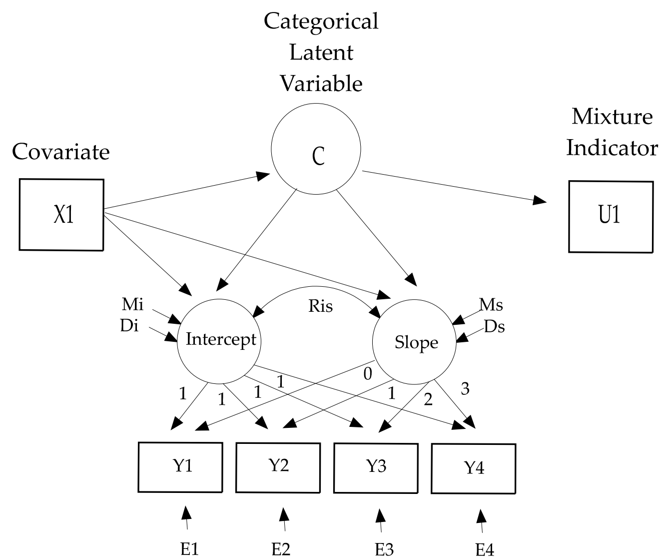

FIG. 10. Effects of the categorical latent variable (C) on the mixture indicator (U).

Model fitting procedures for the model depicted in Figure 10 resulted in a log likelihood H0 value of -1395.05, an entropy estimate of .856, and a BIC value of 3003.124.

Parameter estimates for the regression of U on C varied across latent class. The thresholds for class 1, -3.183, $t = -3.967$, $p < .01$, and class 3, 1.469, $t = 3.920$, $p < .01$ were significant. The threshold for class 2 was not significant, -0.536, $t = -1.167$, $p > .05$.

The mixture indicator U1 is a dichotomous response variable with outcomes 0 and 1 (0 = no alcohol use and 1 = alcohol use at T5). The odds ratio is expressed as the ratio of the odds for those with U1 = 1 to the odds for those with U1 = 0. The parameters (given above) associated with U1 express the log change in log odds from U1 = 0 to U1 = 1. Here, the odds ratio indicates how the odds of the latent class variable, C, changes as U1 changes from U1 = 0 to U1 = 1. In our example the number of categorical

INPUT 6

MPlus Input Statements for the Effects of the Latent Classes on the Mixture Indicator

```
Title: Mixture modeling: 3 latent classes U on C
Data: File is mixexample.dat;
Variable: Names are id x1 y1 y2 y3 y4 u1;
Usevariables are x1 y1 y2 y3 y4 u1;
Classes = c(3);
categorical=u1;
Analysis: Type = mixture;
          miteration = 300;
Model:  %overall%
          int by y1-y4@1;
          slp by y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0];
          y1 y2 y3 y4;
          int*.109 slp*0;
          int with slp*0;
          int on x1;
          slp on x1;
          [c#1*0];
          c#1 on x1;
          c#2 on x1;
          %c#1%
          [int*4.096 slp*.843];
          y1 y2 y3 y4;
          int*1.986 slp*.160;
          int with slp*-.531;
          int on x1;
          slp on x1;
          ! effect of latent class on mixture indicator [binary distal outcome]
          [u1$1*1];
          %c#2%
          [int*2.758 slp*.516];
          y1 y2 y3 y4;
          int*1.418 slp*.240;
          int with slp*-.593;
          int on x1;
          slp on x1;
          ! effect of latent class on mixture indicator [binary distal outcome]
          [u1$1*.5];
          %c#3%
          [int*1.233 slp*0];
          int*1.018 slp*.140;
          int with slp*-.193;
          int on x1;
          slp on x1;
          ! effect of latent class on mixture indicator [binary distal outcome]
          [u1$1*0];
Output: tech1 tech7;
```

latent classes is 3 and the number of levels in the mixture indicator is 2. The last latent class variable (c = 3) serves as the reference category and is fixed at zero. Hence the log odds for latent class 1 is

$$g(c = 1) = \beta_0 + \beta_1(U1 = 1) + \beta_2(U1 = 0) = \beta_0 + \beta_1, \tag{12}$$

and the log odds for latent class 2 is

$$g(c = 2) = \beta_0 + \beta_1(U1 = 0) + \beta_2(U1 = 1) = \beta_0 + \beta_2. \tag{13}$$

The log odds ratio comparing two latent classes, for example, latent class 1 versus latent class 3, is *log(g(latent class 1, latent class 3)) = g(latent class 1) - g(latent class 3)* = $\beta_0 + \beta_1$.

Table 1 shows the resulting estimates of the relation between class membership and subsequent alcohol use. While the normative class (Class 3) has a probability of .19 of engaging in alcohol use at T5, the two comparison classes have elevated probabilities of .63 and .96, respectively. Table 1 also gives the corresponding odds and odds ratios when comparing a class with the normative class.

Table 1
Probability of Subsequent Alcohol Use Given Class Membership

	PROB	ODDS	RATIO
Class 1	.96	24.12	104.79
Class 2	.63	1.71	7.43
Class 3	.19	.23	1.00

The logistic regression coefficients for the effects of X1 on classes 1 and 2 were 3.328, t = 5.763, p < .01, and 1.814, t = 2.801, p < .01, respectively. The logistic regression coefficient for the last class was fixed at zero to serve as the reference category.

The logistic model has similar characteristics with a more general class of linear models. In this model, a function of the mean of the response variable is assumed to be linearly related (with link function) to the independent (X) variables. There is a link between the random and systematic components of the dependent variable in the logistic model.

Logit estimates are used to obtain a predicted probability ($\hat{p}$) of an event. In Mplus, this is expressed as $p = \frac{1}{1+e^{-z}}$, where z = logit(p) $\equiv \log(\frac{p}{1-p}) = \beta_0 + \beta_1{}' X$ and β_0 and β_1 are intercept and slope, respectively.

Intercepts for the latent classes were estimated at values of -10.599, t = -5.775, p < .01, and -5.765, t = -2.924, p < .01, for classes 1 and 2, respectively. The intercept for the last class was fixed at zero to serve as the reference category.

For class 1, the estimate for the intercept, *Mi* = -1.524, t = -1.171, p > .05, was not significant whereas the estimate for the slope, *Ms* = 2.332, t = 5. 191, p < .001, was significant. Estimates for the latent intercept, *Ds* = 1.592, t = 2.695, p < .01, and slope, *Di* = .120, t = 2.369, p < .05, variance also were significant. The covariance between the intercept and slope factors was significant, *R* = -0.394, t = -2.966, p < .01.

For class 2, estimates for the intercept, *Mi* = 1.033, t = .690, p > .05, and slope, *Ms* = 0.775, t = 0.861, p > .05, means were not significant. Estimates for the latent intercept, *Di* = 1.322, t = 1.461, p > .05, and slope, *Ds* = .241, t = 1.533, p > .05, variance also were not significant. The covariance between the intercept and slope factors was not significant, *R* = -.592, t = -1.630, p> .05.

For class 3, the estimates for the intercept, *Mi* = .431, t = 1.142, p > .05, and slope, *Ms* = .560, t = 1.881, p > .05, means were not significant. The estimate for the latent intercept variance, *Di* = .093, t = 2.273, p < .05, was significant whereas the slope variance estimate, *Ds* = .058, t = 1.916, p < .05, was not significant. The covariance between the intercept and slope factors was not significant, *R* = -.003, t = -.114, p > .05.

For class 1, the regression coefficients for the effect of *X1* (Figure 9) on the intercept, 1.494, t = 4.134, p < .01, and slope, -.394, t = -3.595, p < .01, were significant. For class 2, the regression coefficients for the intercept, .541, t = 1.171, p > .05, and slope, -.032, t = -.117, p > .05, were not significant. For class 3, the regression estimate for the intercept, .275, t = 2.103, p < .05, was significant whereas the estimate for the slope, -.137, t = -1.376, p > .05, was not.

Figure 11 presents the conditional growth mixture model which includes the effects of the covariate, X1, on the mixture indicator, U1. This effect indicates that, for each class, the probabilities of U vary as a function of X. The U variables are statistically independent for each C class. Input specifications for the effects of the covariate on the mixture indicator are presented in Input 7. Key specifications of the regression of the mixture indicator on the covariate are given by the input statement u1 on x1*;. Note that the specification of the regressions of U on X are specified as varying parameters in each of the latent classes and appear in each of the latent class specifications.

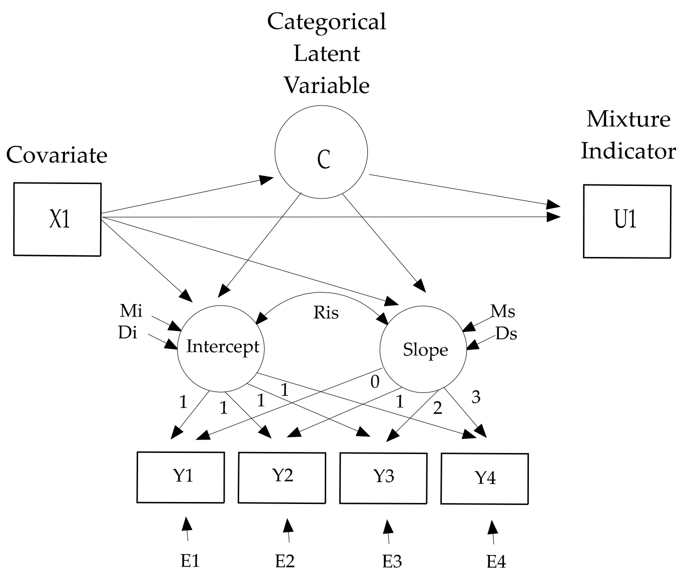

FIG. 11. Effects of the covariate (X1) on the mixture indicator (U1).

Model fitting procedures for the model depicted in Figure 11 resulted in a log likelihood H0 value of -1392.61, an entropy estimate of .859, and a BIC value of 3033.42.

Thresholds for the regression of U on X were 2.927, -304, and -1.242 for class 1, class 2, and class 3, respectively. For each class, the effects were not significant. For classes 1, 2, and 3, the thresholds for the regression of U on C were not significant, 7.043, $t = 1.060$, $p > .05$, -1.596, $t = -.537$, p > .05, and -2.114, $t = -.926$, $p > .01$, respectively.

INPUT 7

MPlus Input Statements for the Effects of the Covariate on the Mixture Indicator

```
Title: Mixture modeling: 3 latent classes U on X
Data: File is mixexample.dat;
Variable: Names are id x1 y1 y2 y3 y4 u1;
Usevariables are x1 y1 y2 y3 y4 u1;
Classes = c(3);
categorical=u1;
Analysis: Type = mixture;
          miteration = 300;
Model:    %overall%
          int by y1-y4@1;
          slp by y1@0 y2@1 y3@2 y4@3;
          [y1-y4@0];
          y1 y2 y3 y4;
          int*.109 slp*0;
          int with slp*0;
           int on x1;
          slp on x1;
          [c#1*0];
          c#1 on x1;
          c#2 on x1;
          %c#1%
          [int*4.096 slp*.843];
          y1 y2 y3 y4;
          int*1.986 slp*.160;
          int with slp*-.531;
          int on x1;
          slp on x1;
          [u1$1*1];
          ! effect of covariate (x1) on mixture indicator (u1)
          u1 on x1*;
          %c#2%
          [int*2.758 slp*.516];
          y1 y2 y3 y4;
          int*1.418 slp*.240;
          int with slp*-.593;
           int on x1;
          slp on x1;
          [u1$1*.5];
          ! effect of covariate (x1) on mixture indicator (u1)
          u1 on x1*;
          %c#3%
          [int*1.233 slp*0];
          int*1.018 slp*.140;
          int with slp*-.193;
          int on x1;
          slp on x1;
          [u1$1*0];
          ! effect of covariate (x1) on mixture indicator (u1)
          u1 on x1*;
Output: tech1 tech7;
```

The logistic regression coefficients for the effects of X1 on classes 1 and 2 were 3.137, $t = 5.720$, $p < .01$, and 1.902, $t = 2.913$, $p < .01$, respectively. Intercepts for the latent classes 1 and 2 were estimated at values of -10.060, $t = -5.738$, $p < .01$ and -6.099, $t = -3.096$, $p < .01$, respectively.

For class 1, the estimate for the intercept, $Mi = -1.692$, $t = -1.383$, $p > .05$, was not significant whereas the estimate for the slope, $Ms = 2.354$, $t = 5.671$, $p < .001$, was significant. Estimates for the latent intercept, $Ds = 1.567$, $t = 2.638$, $p < .01$, and slope, $Di = .113$, $t = 2.154$, $p < .05$, variances also were significant. The covariance between the intercept and slope factors was significant, $R = -0.384$, $t = -2.982$, $p < .01$.

For class 2, estimates for the intercept $Mi = 1.087$, $t = .714$, $p > .05$, and slope means, $Ms = 0.543$, $t = .680$, $p > .05$, were not significant. Estimates for the latent intercept, $Di = 1.316$, $t = 1.415$, $p > .05$, and slope, $Ds = .209$, $t = 1.298$, $p > .05$, variances also were not significant. The covariance between the intercept and slope factors was not significant, $R = -.573$, $t = -1.560$, $p > .05$.

For class 3, estimates for the intercept, $Mi = .431$, $t = 1.149$, $p > .05$, and slope, $Ms = .561$, $t = 1.735$, $p > .05$, means were not significant. The estimates for the latent intercept, $Di = .095$, $t = 2.246$, $p < .05$, and slope , $Ds = .058$, $t = 1.706$, $p < .05$, variances, however, were significant. The covariance between the intercept and slope factors was not significant, $R = -.002$, $t = -.088$, $p > .05$.

For class 1, the regression coefficients for the effect of the covariate, *X1*, on the growth parameters (Figure 9), 1.494, $t = 4.134$, $p < .01$, and -.394 $t = -3.595$, $p < .01$, were significant for the intercept and slope, respectively. For class 2, the coefficients for intercept and slope were not significant, .541, $t = 1.171$, $p > .05$ and -.032, $t = -.117$, $p > .05$, respectively. For class 3, the estimate for the intercept, .275, $t = 2.103$, $p < .05$, was significant whereas the estimate for the slope, -.137, $t = -1.376$, $p > .05$, was not significant.

Model fitting statistics for the various competing mixture models and parameter estimates for the growth factors are presented in Appendix 1. Parameter estimates for regression models and percentage of individuals in each class are presented in Appendix 2.

DISCUSSION

The models proposed in this chapter clearly necessitate analyses specifically designed for longitudinal data. Moreover, to study development in a comprehensive fashion, one must consider change and variability over time. An underlying assumption of the more conventional methods for the analysis of change is that the data come from a single population (i.e., all individuals belong to one and the same population) that encompasses varying trajectories. In particular, these methods assume that covariates have the same influence on the growth factors for all individual trajectories in the population. However, individuals with different growth/change trajectories may not only have different growth shapes, but also differing antecedents (or pre-transitional characteristics), different concurrent processes, and different consequences. A growth mixture modeling approach, therefore, enables the researcher to study qualitatively different developmental processes across individuals belonging to several, presumably identifiable, subgroups.

In this paper, we presented the specifications necessary for conducting latent variable growth mixture modeling. This type of modeling includes models with a combination of categorical and continuous latent variables. A categorical latent variable is used to represent a mixture of subpopulations in which membership is not known but inferred from the data. In statistics, this is referred to as finite mixture modeling. Latent variable mixture modeling consists of two parts: the mixture portion of the model and the structural portion of the model. In the mixture portion of the model, the categorical latent variable is allowed to influence binary and/or ordered categorical observed outcomes referred to as latent class indicators, or continuous observed outcomes, and to be regressed on observed variables or covariates. In growth mixture modeling, the modeling of repeated measures of the binary or ordered categorical latent class indicators also is possible. In the structural portion of the model, all parameters are allowed to vary across the latent classes. The continuous latent variables are allowed to influence continuous observed outcomes, but not the categorical observed outcomes. Analyses are carried out using a ML estimator.

Limitations of the Mixture Approach

It was shown that growth mixture modeling can be used instead of the more traditional latent growth curve model to reduce dimensionality and estimate a separate regression model for each subpopulation (group). One potential drawback of this approach is that there is no guarantee of model convergence. Some computer programs, such as Mplus (Muthén & Muthén, 1998) and Latent GOLD (Vermunt & Magidson, 2000), employ Fisher Scoring, Newton Raphson, or EM algorithms to aid in the convergence to a local maxima. Moreover, it is possible that even when convergence is achieved, multiple solutions may be found using different start values. Software developers could increase the likelihood of convergence and the likelihood of a unique model solution if they incorporated procedures for including random starting values.

The latent class model also assumes that the same measurement and structural model holds for all latent classes, although this assumption may not always be tenable in practice. In large datasets, a large number of latent classes could be required to adequately explain the variation that exists across all individuals in the sample. Because latent class analysis uses a multinomial logit for determining differences between classes , this approach ignores interindividual differences that exist within classes. That is, response probabilities for all individuals of a given class are identical. Therefore, as shown in Inputs 4 - 7, structural relationships between the latent (continuous and binary) and observed (covariate) variables should be considered, as their inclusion may help account for some of the interindividual differences that exist within the latent classes.

Growth mixture modeling is an important new development in the study of change. With new modeling opportunities come the potential for exploring new and more complex theories of development in a plethora of behavioral fields of study. The growth mixture modeling approach presented here is strengthened by its association with the general latent variable modeling framework.

REFERENCES

Arminger, G., & Stein, P. (1997). Finite mixtures f covariance structure models with regressors. Sociological Method & Research, 26, 148-182.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade, & P. B. Baltes (Eds.), Longitudinal research in the study of behavior and development (pp. 1-40). New York: Academic Press.

Collins, L. (1991). The measurement of dynamic latent variables constructs in longitudinal aging research: quantifying adult development. Experimental Aging Research, 17, 13-20.

Collins, L., & Cliff, N. (1990). Using the longitudinal Guttman simplex as a basis for measuring growth. Psychological Bulletin, 108, 128-134.

Collins, L., Graham, J. W., Long, J., & Hansen, W. B. (1994). Crossvalidation of latent class models of early substance use onset. Multivariate Behavioral Research, 29, 165-183.

Collins, L., Graham, J. W., Rousculp, S. S., & Hansen, W. (1997). Heavy caffeine use and the beginning of the substance use onset process: An illustration of latent transition analysis. In K. J. Bryant, M. Windle, and S. G. (Eds.), The science prevention: Methodological advances from alcohol and substance abuse research (pp. 79-99). Washington D. C.: American Psychological Association.

Collins, L., & Horn, J. L. (1991). Best methods for the analysis of change. Washington, DC: APA Press.

Collins, L., & Sayer, A. (2001). New methods for the analysis of change. American Psychological Association, Washington D.C.

Collins, L., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variable. Multivariate Behavioral Research, 27, 131-157.

Clogg, C. C. (1995).Latent class models. In G. Arminger, C. C. Clogg & M. E. Sobel (Eds.), Handbook of statistical modeling for the social and behavioral sciences (pp. 311-359). New York: Plenum Press.

Dolan, C. V., & van der Maas, H. L. J. (1998). Fitting multivariate normal finite mixtures subject to structural equation modeling. Psychometrika, 63, 227-253.

Duncan, S. C., & Duncan, T. E. (1996). A multivariate latent growth curve analysis of adolescent substance use. Structural Equation Modeling, 4, 323-347.

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). An Introduction to Latent Variable Growth Curve Modeling: Concepts, Issues, and Applications. Mahwah NJ: Lawrence Erlbaum Associates.

Everitt, B. S., & Hand, D. J. (1981). Finite mixture distributions. London: Chapman and Hall.

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

Graham, J. W., Collins, L. M., Wugalter, S. E., Chung, N. K., & Hansen, W. B (1991). Modeling transitions in latent stage-sequential processes: A substance use prevention example. Journal of Consulting and Clinical Psychology, 59, 48-57.

Hansen, W.B. & Graham, J.W. (1991). Preventing alcohol, marijuana, and cigarette use among adolescents: peer pressure resistence training versus establishing conservative norms. Preventive Medicine, 20, 414-430.

Heinen, T. (1996). Latent class and discrete latent trait models: Similarities and differences. Thousand Oaks, CA: Sage Publications

Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. Marketing Science, 16, 39-59.

Langeheine, R., & Rost, J. (1988). Latent trait and latent class models. New York: Plenum Press.

McCutcheon, A. L. (1987). Latent class analysis. Newbury Park, CA: Sage.

Meiser, T., & Ohrt, B. (1996). Modeling structure and chance in transitions: Mixed latent partial Markov-Chain models. Journal of Educational and Behavioral Statistics, 21, 91-109.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. Psychometrika, 55, 107-122.

Muthén, B. O. (2001). Second-generation structural equation modeling with combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In A. Collins L. & Sayer A. (Eds.), New methods for the analysis for change (pp. 291-322) Washington D.C.: American Psychological Association.

Muthén, B., Brown, C. H., Khoo, S., Yang, C., & Jo, B. (1998). General growth mixture modeling of latent trajectory classes: Perspectives and prospects. Paper presented at the meeting of the Prevention Science and Methodology Group, Tempe, AZ.

Muthén, L. K., & Muthén, B. (1998). Mplus: User's guide. Los Angeles, CA: Muthén & Muthén.

Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. Biometrics, 55, 463-469.

Nagin, D. S. (1999). Analyzing developmental trajectories: A semiparametric, group-based approach. Psychological Methods, 4, 139-157.

Ramaswamy, V., DeSarbo, W., Reibstein, D., & Robinson, W. (1993). An empirical pooling approach for estimating marketing mix elasticities with PIMS data. Marketing Science, 12, 103-124.

Tisak, J., & Meredith, W. (1990). Descriptive and associative developmental models. In A. von Eye (Ed.), New Statistical Methods in Developmental Research. New York: Academic Press.

Titterington, D. M., Smith, A. F. M., & Makov, U. E. (1985). Statistical analysis of finite mixture distributions. Chichester, U. K.: John Willey & Sons.

Vermunt, J.K. & Magidson, J. (2000). Latent GOLD 2.0 User's Guide. Belmont, MA: Statistical Innovations Inc.

Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. Psychological Bulletin, 116, 363-381.

Yang, C. C. (1998). Finite mixture model selection with psychometric applications. Unpublished doctoral dissertation. University of California, Los Angeles.

Yung, Y. F. (1997). Finite mixtures in confirmatory factor-analysis models. Psychometrika, 62, 297-330.

APPENDIX 1

MODEL FITTING STATISTICS AND GROWTH FUNCTION ESTIMATES

				Class 1		Class 2		Class 3	
Model	Log H_0	BIC	Ent.	M_i D_i	M_s D_s R_{is}	M_i D_i	M_s D_s R_{is}	M_i D_i	M_s D_s R_{is}
Input 1	-1547.07	3142.09		2.850^{**} 2.607^{**}	0.558^{**} 0.174^{**} -0.021				
Input 2	-1496.36	3093.88	0.796	4.442^{**} 1.278^{*}	0.720^{**} 0.174^{**} -0.390^{*}	1.881^{**} 1.081^{*}	$.484^{**}$ $.232^{**}$ -.134		
Input 3	-1361.26	2877.02	0.847	4.096^{**} 1.986^{**}	0.843^{**} 0.160^{**} $-.531^{**}$	2.758^{**} 1.418^{*}	$.516^{*}$ $.240^{*}$ $-.59^{*}$	1.23^{**} $.109^{*}$	0.180 0.053 -.014
Input 4 (With X1)	-1339.04	2864.56	0.829	-2.468 1.131^{*}	2.526^{**} 0.130^{*} -0.345^{*}	0.479 1.130	0142 0.214 -.497	0.451 0.116	-.083 .068 -.015
Input 5 (C on X1)	-1306.42	2809.97	0.860	-1.778 1.290^{*}	2.421^{**} 0.131^{*} $-.381^{**}$	1.553 1.280	0.254 $.245^{*}$ $-.58^{*}$	0.465 $.089^{*}$	0.602 $.058^{*}$ -.004
Input 6 (C on X1 and U1 on C)	-1395.05	3003.12	0.856	-1.524 1.592^{**}	2.332^{**} 0.120^{**} $-.394^{**}$	1.033 1.322	0.775 0.241 -.592	0.431 $.093^{*}$	0.560 0.058 -.003
Input 7 (C on X1 , U1 on C, and U1 on X1)	-1392.16	3013.42	0.859	-1.692 1.567^{**}	2.354^{**} 0.113^{*} $-.384^{**}$	1.087 1.316	0.543 0.209 -.573	0.431 0.095	0.561 0.058 -.002

$^{*}P<.05$, $^{**}P<.01$

APPENDIX 2

PARAMETER ESTIMATES FOR REGRESSION MODELS AND CLASSIFICATION RATES

		Input 1	Input 2	Input 3	Input 4	Input 5	Input .6	Input 7
Int on X1 Slp on X1	Class-1				1.794** -0.459**	1.576** -0.420**	1.451** -0.389**	1.494** -.394**
	Class-2				0.659 0.124	0.359 0.082	0.546 -0.091	0.541 -0.032
	Class-3				0.266 0.089	0.262 -0.137	2.075 * -0.122	0.275* -0.122
C on X1	Class-1					-10.307** 3.185**	-10.599** 3.282**	-10.1** 3.137**
	Class-2					-6.778 * 2.158 *	-5.765** 1.814**	-6.10** 1.902**
	Class-3							
U1 on C (thresholds)	Class-1						-3.183**	7.043
	Class-2						-0.536	-1.596
	Class-3						1.469**	-2.114
U1 on X1	Class-1							2.927
	Class-2							-0.304
	Class-3							-1.242
% of individuals in each class	Class-1	1.00	0.38	0.40	0.38	0.41	0.43	0.45
	Class-2		0.62	0.30	0.32	0.29	0.26	0.24
	Class-3			0.30	0.30	0.30	0.31	0.31

*P<.05, **P<.01