

## Second-Generation Structural Equation Modeling With a Combination of Categorical and Continuous Latent Variables

### New Opportunities for Latent Class—Latent Growth Modeling

---

Bengt Muthén

This chapter focuses on the interplay between growth curve modeling and structural equation modeling (SEM). Growth curve modeling is concerned with the study of individual differences in development over time, typically captured by random effects, that is, growth parameters that vary across individuals. SEM is concerned with relationships among observed and latent variables. The two areas connect because random coefficients may be viewed as continuous latent variables (see, e.g., Meredith & Tisak, 1990). Growth modeling in a latent variable SEM framework thereby benefits from the full generality of such a framework. For example, in the SEM framework it is convenient to study regressions among the random effects; to do growth analysis in multiple populations; to analyze growth in latent variable constructs measured by multiple indicators; to analyze both parallel and sequential growth processes simultaneously; and to include other model parts that relate to the growth

---

This research was supported by Grant K02 AA 00230 from the National Institute on Alcohol Abuse and Alcoholism (NIAAA), by Grant R21 AA10948 from NIAAA, and by Grant 40859 from the National Institute of Mental Health. The work benefited from discussions in Hendricks Brown's Prevention Science Methodology group and Muthén's Methods Mentoring Meeting. I thank Tom Harford for suggestions regarding the National Longitudinal Survey of Youth application; Hendricks Brown and Jason Liao for discussion related to intervention studies; Linda Muthén for useful comments; and Siek-Toon Khoo, Booil Jo, Noah Yang, Christy Kim Boscardin, and Ching Yu for helpful research assistance.

Muthén, B. (2001). Second-generation structural equation modeling with a combination of categorical and continuous latent variables: New opportunities for latent class/latent growth modeling. In Collins, L.M. & Sayer, A. (Eds.), *New Methods for the Analysis of Change* (pp. 291-322).

Washington, D.C.: APA. (#82)

model, such as mediators and distal outcomes. For an overview of examples of this kind, see B. Muthén and Curran (1997).

The realization that growth curve modeling is possible using a latent variable conceptualization has enriched SEM research in recent years. By its general modeling framework, however, SEM has also enriched growth modeling. SEM is now about to take a radical leap forward in terms of modeling capabilities, and this will further benefit growth curve modeling. As I describe in this chapter, conventional SEM models using continuous latent variables are for the first time integrated with models using categorical latent variables. Although categorical latent variable modeling has long traditions in the form of latent class analysis (LCA), latent profile analysis (LPA), and latent transition analysis (LTA), these traditions have been completely separated from SEM. Categorical latent variable modeling also encompasses finite mixture modeling in general and goes well beyond LCA, LPA, or LTA. New research reviewed here provides an integration of categorical and continuous latent variable models. Given its generality, it is fitting to describe the emerging methodology as second-generation SEM, where the focus is on the generality of latent variable modeling (LVM). This LVM development promises to be extremely beneficial to growth modeling. The aim of this chapter is to briefly introduce new LVM analyses in the form of general growth mixture modeling (GGMM) and to show examples of the new analysis opportunities for growth modeling that are opened up. In this chapter, I discuss several different types of GGMM applications and analyze two different examples. My presentation is as nontechnical as possible to reach applied researchers. The analyses of the examples were carried out by the new computer program Mplus (L. K. Muthén and Muthén, 1998). Input specifications are available from the website <http://www.statmodel.com>.

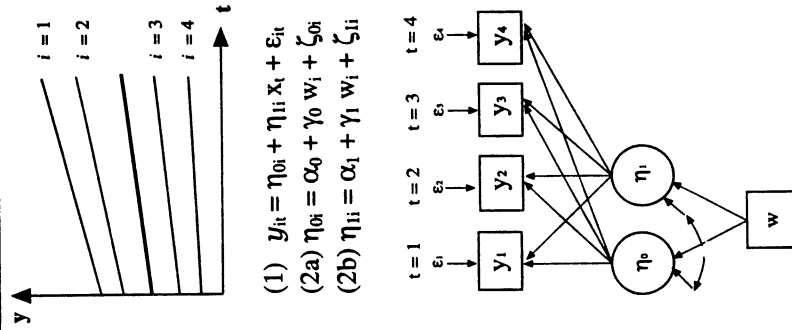
### Conventional Latent Growth Modeling

Conventional growth modeling with random coefficients is a useful starting point for introducing the new modeling ideas. The basic ideas, scope, and limitations of conventional growth modeling are discussed here.

Figure 10.1 shows three ways of representing growth modeling: (a) using a graph of individual trajectories, (b) using a two-level model, and (c) using a latent variable model. The top panel of Figure 10.1 shows trajectories for 4 individuals. The mean trajectory is also given as a solid line. One can see that there is variation in the initial status as well as the growth rate. Individuals who start higher grow faster.

The middle panel of Figure 10.1 translates the individual-differences idea into a random-effect model expressed as a multilevel model. The Level 1 equation describes variation as a function of the time score  $x_i$  at time point  $t$ . For

FIGURE 10.1  
Three representations of individual differences in growth.



example, in an educational study  $t$  can be Grade 7, Grade 8, and Grade 9, and the  $x_t$ s can be scores representing linear growth—for example, 0, 1, 2—or nonlinear growth—for example, 0, 1, 2.5. Here,  $\eta_0$  and  $\eta_1$  represent intercept and slope coefficients in the regression of  $y$  on  $x$ . The individual differences are captured by letting the intercept and slope vary across individuals, so that the  $\eta$  coefficients have subscript  $i$ . These coefficients are referred to as *random coefficients* or *random effects*. Although this allows for across-individual heterogeneity, all individuals are assumed to come from one and the same population so that the analysis benefits from estimating all individuals together. The two Level 2 equations describe variation across individuals for the intercepts and slopes and relate this variation to a background variable  $w$ , referred to as a *time-invariant covariate* (time-varying covariates can be included in the Level 1 equation). The fact that the individuals come from a single population is re-

flected by the absence of a subscript  $i$  for the Level 2 parameters of  $\alpha$  and  $\gamma$ . Further single population parameters appear in the covariance matrix for the residuals  $\zeta$  and for the residuals  $\epsilon$ . Inserting the intercept and slope of the Level 2 equations into the Level 1 equation gives the model in a form referred to as the *mixed linear model*.

The bottom panel of Figure 10.1 translates the two-level model into a latent variable framework in which the random effects are reconceptualized as latent variables. The Level 1 equation and the Level 2 equations correspond to a measurement and a structural part of a latent variable model, respectively. An advantage of growth modeling in a latent variable framework is that the framework opens up possibilities for general structural equation modeling, including regressions among random coefficients; multiple population analysis with flexible cross-group differences in covariance matrix structures; growth modeling of latent variable constructs with multiple indicators; analysis of multiple processes; and the inclusion of other model parts, such as mediational modeling and the analysis of distal outcomes.

### A Simple Conventional Growth Model and Its Limitations

Consider a simple linear growth model with an outcome  $y_{it}$  for individual  $i$  observed at time point  $t$ ,

$$y_{it} = \eta_{0i} + \eta_{1i}x_t + \epsilon_{it}, \quad (10.1)$$

where  $\eta_{0i}$  is the intercept factor,  $\eta_{1i}$  is the growth rate factor,  $x_t$  is a time score, and  $\epsilon_{it}$  is the time-specific residual assumed to be normally distributed. A given individual  $i$  has the values  $\eta_{0i}$  and  $\eta_{1i}$  on the two growth factors. The growth process for this individual develops over time as  $t$  changes as

$$\eta_{0i} + \eta_{1i}x_t. \quad (10.2)$$

This is individual  $i$ 's trajectory, describing the systematic part of the variation of the outcome at different time points. The individual's outcome at a certain time point  $t$ ,  $y_{it}$ , is equal to the sum of the systematic part of the variation plus the time-specific residual  $\epsilon_{it}$ .

As mentioned above, Equation 10.1 is often referred to as the *Level 1 equation*, describing the repeated measures over time. The Level 2 equation describes the variation in the  $\eta_{0i}$  and  $\eta_{1i}$  factors as a function of covariates  $x$ ,

$$\eta_{0i} = \alpha_0 + \sum_r \gamma_{0r}w_{ir} + \zeta_{0i}, \quad (10.3)$$

$$\eta_{1i} = \alpha_1 + \sum_r \gamma_{1r}w_{ir} + \zeta_{1i}, \quad (10.4)$$

where  $\alpha$  coefficients are intercept parameters,  $\gamma$  coefficients are regression

weights for the covariates  $w_r$ , and the  $\zeta$ s represent residuals assumed to be bivariate normally distributed.

The parameters of Equations 10.1, 10.3, and 10.4 can be estimated by maximum likelihood. Given such estimates, individual estimates of the growth factor values  $\eta_{0i}$  and  $\eta_{1i}$  can be obtained by the Bayesian approach of maximizing the posterior distribution of the factors given the individual's observed data. In statistics this is referred to as *empirical Bayes estimation*, whereas in psychometrics it is referred to as *factor score estimation using the regression method*.

The five trajectories shown in Figure 10.1 follow the random coefficient growth model of Equations 10.1, 10.3, and 10.4. The bold line is the average trajectory, which is evaluated at the means of the two growth factors  $\eta_{0i}$  and  $\eta_{1i}$ ; that is, these factor means are inserted into Equation 10.2. In addition, trajectories are shown for intercept growth factor values of 1 and 2 SD below and above the mean, with corresponding slope factor values derived from the bivariate normal distribution. The variation across these five trajectories is due to the individual-specific influence of the covariate  $x_{it}$  and the residuals  $\zeta_{0i}$  and  $\zeta_{1i}$ .

Some of the trajectories in Figure 10.1 represent quite different development over time. For example, if the outcome represents reading skills in Grade 1, the two bottom lines labeled  $i = 3$  and  $i = 4$  may correspond to somewhat problematic or very problematic reading development, whereas the two top lines labeled  $i = 2$  and  $i = 1$  may correspond to good or excellent reading development. An assumption of the conventional growth model is that all individuals belong to one and the same population. In particular, it is assumed that the  $w$  covariates have the same influence on the growth factors for all trajectories; that is, the  $\gamma$  covariate slopes of Equations 10.3 and 10.4 are the same for all individuals. Using the example of the two types of good reading development versus the two types of poor reading development, this assumption of a homogeneous population may not be realistic. For example, the variation in reading development among poor readers may be influenced more by the school environment, whereas the variation in reading development for good readers may be influenced more by the home environment or vice versa. In addition, the variances of the residuals may differ for these two types of development, violating a second assumption in the conventional growth model. Although the conventional growth model captures individual differences in trajectories, it is not always realistic to assume that a single-population model can account for all types of individual differences.

In general, the two types of development just discussed may have not only different antecedents but also different growth shapes, different concurrent processes, and different consequences. For example, problematic first-grade reading development may have a nonlinear shape, may co-occur with the development

of aggressive classroom behavior, and may increase the probability of subsequent deficits in achievement development or school dropout. This type of developmental heterogeneity is presumably quite common. In mental health, drug, and alcohol research the recognition of heterogeneity has led to theories of multiple developmental pathways (see, e.g., Moffitt, 1993; on adolescent-limited vs. life-course-persistent antisocial behavior), subtypes (see, e.g., Zucker, 1994, and Schulenberg, O'Malley, Bachman, Wadsworth, & Johnston, 1996, on alcoholism; and Nagin, Farrington, & Moffitt, 1995, on criminal offenders), and different disease processes (see, e.g., Pearson, Morrell, Landis, Carter, & Brant, 1994, on prostate cancer development). Pearson et al.'s (1994) study illustrates some key modeling ideas. It examined the development of prostate-specific antigen (PSA) and related this to later occurrence of tumors and cancers. Whereas the normative development of PSA over age can be described by a linear growth model with variation in intercept and slope, development leading to tumors and cancers is characterized by a change from linear to exponential PSA growth in which the severity of the outcome corresponds to different forms of the growth curves.

In summary, the conventional growth model allows heterogeneity corresponding to different growth trajectories across individuals and captures that by variation in the continuous growth factors. However, the conventional growth model cannot capture heterogeneity that corresponds to qualitatively different development. Although multiple-group SEM growth modeling is a flexible tool for studying qualitatively different development across individuals belonging to known groups, the problem here is that the group membership is typically not known but needs to be inferred from the data. In the next section, I discuss new methodology that introduces categorical latent variable modeling to greatly add to the capability of capturing heterogeneity in development.

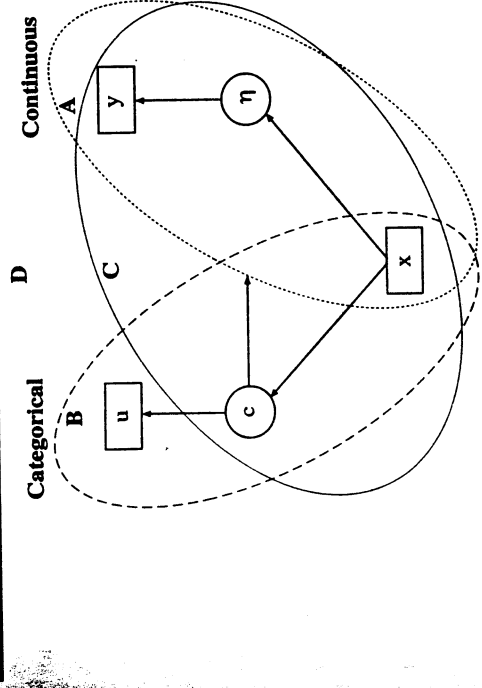
## Second-Generation Structural Equation Modeling

Second-generation SEM uses a combination of categorical and continuous latent variables. The ideas can be described with the help of the prototypical model structure of Figure 10.2. The figure shows four different modeling frameworks, labeled A, B, C, and D. After describing the statistical techniques, I discuss each framework in turn, with an emphasis on C and D, which provide the new growth modeling opportunities studied in the remainder of this chapter.

What follows is a brief technical description of the model in Framework D, its estimation, and testing. A nontechnical reader can skip this section because the understanding of later sections does not depend on it. The full specifications of the general LVM are given in the Mplus User's Guide (L. K. Muthén & Muthén, 1998), which generalizes the approach of B. Muthén and Shedden

FIGURE 10.2

### A general framework for latent variable modeling.



(1999). The formulas below are for a simple example. With linear growth in latent class  $k$  ( $k = 1, 2, \dots, K$ ) the Level 1 growth part of the model is the measurement model,

$$y_{itk} = \eta_{0ik}x_i + \eta_{1ik}x_i + \kappa_{ik}w_{itk} + \epsilon_{itk}, \quad (10.5)$$

whereas the Level 2 part of the model is the structural model,

$$\eta_{0ik} = \alpha_{0k} + \gamma_{0k}w_{itk} + \zeta_{0ik}, \quad (10.6)$$

$$\eta_{1ik} = \alpha_{1k} + \gamma_{1k}w_{itk} + \zeta_{1ik}, \quad (10.7)$$

where  $w_{itk}$  is a time-varying covariate and  $w_{itk}$  is a time-invariant covariate.

A second part of the model includes predictors of the latent class membership variable vector  $c_i$  of length  $K$ , where  $c_{ik} = 1$  if individual  $i$  belongs to class  $k$  and 0 otherwise. For example, with a single predictor  $z_{ik}$ ,

$$P(c_{ik} = 1 | z_{ik}) = \frac{e^{\alpha_k + \gamma_k z_{ik}}}{\sum_{k=1}^K e^{\alpha_k + \gamma_k z_{ik}}}, \quad (10.8)$$

a multinomial logistic regression that uses the standardization of zero coefficients for the last class,  $\alpha_{kK} = 0$ ,  $\gamma_{kK} = 0$  (see also Agresti, 1990). This gives the logit for the odds of class  $k$  relative to class  $K$ ,

$$\text{logit } c = \log[P(c_{ik} = 1 | z_{ik}) / (P(c_{iK} = 1 | z_{ik}))] = \alpha_{ik} + \gamma_{ik}z_{ik}, \quad (10.9)$$

which is a regular logistic regression in the case of two latent classes.

A third part of the model includes categorical indicators of the latent class

variable,  $u$  say. The full model can, however, be identified without such latent class indicators. In line with latent class modeling conditional independence is specified for the  $u$ s given latent class and the background variables of  $x'_i = (z_i, w_i, w_{i1}, w_{i2}, \dots, w_{iT})$ . Assuming binary  $u$ s, a logit regression is specified for each  $u_j$ , ( $j = 1, 2, \dots, p$ ), for example,

$$\text{logit}[P(u_{ij} = 1 | c_{ik} = 1, z_i)] = \lambda_{ijik} + \kappa_{ij} z_{ik} \tag{10.10}$$

This translates into a probability for  $u_j$ ,

$$P(u_{ij} = 1 | c_{ik} = 1, z_i) = \frac{1}{1 + e^{-\lambda_{ijik} + \kappa_{ij} z_{ik}}} \tag{10.11}$$

which, when there is no covariate  $z$ , gives the usual conditional probability of latent class analysis. If  $\kappa_{ij}$  is different from zero there is a direct effect from the covariate to the latent class indicator, indicating within-class heterogeneity.

Estimation of model parameters uses maximum likelihood and the EM algorithm. Here,  $c$  is considered the missing data in the EM algorithm. EM maximizes the expected complete-data log likelihood  $\log L_c$  for the vector of outcomes at the  $T$  time points  $y'_i = (y_{i1}, y_{i2}, \dots, y_{iT})$  and the vector of  $c$  indicators  $u'_i = (u_{i1}, u_{i2}, \dots, u_{ip})$ , conditional on the vector  $x'_i = (z_i, w_i, w_{i1}, w_{i2}, \dots, w_{iT})$ ,

$$E(\log L_c) = E \left\{ \sum_{i=1}^n [\log(c_i | x_i) + \log(u_i | c_i, x_i) + \log(y_i | c_i, x_i)] \right\}, \tag{10.12}$$

where  $(c_i | x_i)$  denotes the probabilities corresponding to Equation 10.8,  $(u_i | x_i)$  denotes the probabilities of the latent class indicators, and

$$(y_i | c_i, x_i) = [N(\mu_{ik}, \Sigma_k)]^{c_k} \tag{10.13}$$

denotes a mixture of normal distributions. The EM iterations consist of  $E$  steps followed by  $M$  steps, followed by new  $E$  and  $M$  steps until convergence. The  $E$  step computes the expectation of  $\log L_c$ , which centers on the posterior probability of individual  $i$  belonging to class  $k$ ,

$$p_{ik} = P(c_{ik} = 1 | y_i, u_i, x_i) = P(c_{ik} = 1 | z_{ik}) N(\mu_{ik} | c_{ik} = 1) / (y_i, u_i | x_i). \tag{10.14}$$

In this way, the  $E$  step computation of  $p_{ik}$  draws on information from three observed data sources,  $y_i$ ,  $u_i$ , and  $x_i$ . The maximization of Equation 10.12 in the  $M$  step involves three separate maximizations using the posterior probabilities: a multinomial regression optimization for  $c$  on  $x$ , a latent class optimization for the  $u$  indicators, and an optimization of

$$E \left[ \sum_{i=1}^n \log(y_i | c_i, x_i) \right] = \sum_{i=1}^n \sum_{k=1}^K p_{ik} \log N(\mu_{ik}, \Sigma_k). \tag{10.15}$$

The expression in Equation 10.15 corresponds to a simultaneous optimization for the  $K$  classes with sampling weights  $p_{ik}$ . Using SEM language, the growth mixture modeling with respect to the  $y$  part of the model can therefore be seen as a multiple-group analysis but with unobserved group membership. Here, any parameter in  $\mu_{ik}$ ,  $\Sigma_k$  may vary across the  $K$  classes. Parameters may be fixed or constrained to be equal in all three parts of the model. Equalities are not permitted between parameters in the different model parts.

Model testing is somewhat more complex with mixture models. For comparison of fit of models that have the same number of classes and are nested, the usual likelihood-ratio chi-square difference test can be used. Comparison of models with different numbers of classes, however, cannot be made using the likelihood-ratio chi-square test common in SEM. It may instead be accomplished by standard information criteria such as Akaike's and Bayesian (BIC). The measures are based on the negative of the log likelihood of the model, with a penalty for the number of parameters. In this way, smaller values indicate better models. In this chapter, I use BIC. The BIC penalty is the number of parameters  $r$  multiplied by  $\ln n$ , where  $n$  is the sample size (Schwartz, 1978),

$$\text{BIC} = -2 \log L + r \ln n. \tag{10.16}$$

Further technical aspects of LVM and its estimation by maximum likelihood were described by B. Muthén, Shedden, and Spisic (1998). The analyses of all examples presented here were carried out by this method as implemented in the computer program Mplus (L. K. Muthén & Muthén, 1998). Mplus is a program that replaces my LISCOMP program. Despite its ability to carry out such general types of models as indicated by Framework D, Mplus has a simple model specification language without matrices or equations that is suitable for applied researchers.

### Framework A: Conventional Latent Variable Growth Modeling

The modeling framework labeled A in Figure 10.2 has observed continuous outcomes or latent variable indicators  $y$ , latent continuous variables  $\eta$ , and observed background variables  $x$ . This encompasses conventional SEM as it has been practiced in the past 25 years (see, e.g., Bollen, 1989). When the observed outcomes represent repeated measures over time, the latent variables are the growth factors, and the background variables are the covariates as in Equations 10.1, 10.3, and 10.4. Framework A was used by B. Muthén and Curran (1997) in their discussion of extensions of conventional growth modeling, particularly for randomized treatment-control studies.

### Framework B: Latent Class Modeling

The modeling framework labeled B in Figure 10.2 has categorical latent class indicators  $u$ , a categorical latent class variable  $c$ , and observed background variables  $x$ . This includes conventional latent class analysis (see, e.g., Clogg, 1995) and more recent extensions that add covariates (see Bandeen-Roche, Miglioretti, Zeger, & Rathouz, 1997; Dayton & Macready, 1988; Formann, 1992; and van der Heijden, Dessens, & Böckenholt, 1996). When the latent class indicators represent repeated measures over time, LIA (see, e.g., Collins & Wugalter, 1992) specifies a latent class variable for each time point and analyzes the transition probabilities between classes.

### Framework C: Finite Mixture Modeling

The modeling framework labeled C in Figure 10.2 extends Framework A by the categorical latent class variable. The arrow from the latent class variable  $c$  to model framework A indicates that the parameters of A can be different for different latent classes. The analogy to conventional SEM is that the latent classes represent multiple populations or groups, but in contrast to multiple-group SEM, group membership is unobserved. This analogy makes it clear that Framework C provides very flexible modeling given that multiple-group SEM allows group differences in any of the parameters. Model C analysis also produces a counterpart to factor scores for continuous latent variables. Given that the latent class variable is categorical, the factor score notion is generalized to posterior probabilities of membership in the different latent classes. A person may be classified into the class that has the highest probability.

### Mixture Models for Clustering

In statistics, the latent classes are viewed as mixture components or missing data. A key reference for mixture analysis is Titterton, Smith, and Makov (1985). Most of the statistical research using Framework C does not include latent continuous variables. A typical application is a multivariate model for the outcomes in which the mixture components have the same covariance matrix and different mean vectors. This is a form of cluster analysis given that posterior probabilities of class membership are produced (see, e.g., McLachlan & Basford, 1988). An example is the classic three-component analysis of the Fisher Iris data (see, e.g., Everitt & Hand, 1981). In psychometrics, related classic work considers latent profile analysis (see, e.g., Bartholomew, 1987; Gibson, 1959), assuming a diagonal covariance matrix for each component.

### Complier Average Causal Effect Modeling

Little and Yau (1998) gave a novel example of a latent class variable viewed as missing data. This introduces the concept of complier average causal effect (CACE) estimation. In intervention studies, not all individuals invited for treatment actually choose to participate. Those who do choose to participate are typically not a random subset of those invited. Because of the randomization, the control group has the same subgroup of individuals who potentially would participate had they been invited. Here, latent classes corresponding to those who participate versus those who do not participate are used to give an assessment of treatment effects comparing participants in the treatment group with potential participants in the control group. For those randomized into the treatment group, the latent class membership is observed. For individuals in the control group, however, information on category membership is missing, that is, class membership is a latent categorical variable. This is a promising new technique that can also be used in growth modeling contexts.

### Mixture Structural Equation Modeling

Framework C has been considered in an SEM context. For example, Blåfield (1980); Jedidi, Ramaswamy, DeSarbo, and Wedel (1996); Arminger and Stein (1997); Yung (1997); Jedidi, Jagpal, and DeSarbo (1997); and Arminger, Stein, and Wittenberg (1999) have studied mixture confirmatory factor analysis and SEM. The models studied by these authors, however, are more limited than what is shown as Framework C in Figure 10.2. This is because the models do not allow for the regression of the latent class variable on the background variables. This limitation is shared by most mixture modeling, two recent exceptions being those by Little and Yau (1998) and by Nagin (1999). Not including the regression of the latent class variable on the background variables has two drawbacks. First, the probabilities of latent class membership are taken to be the same for individuals with different values on the background variables, which may be unrealistic in many settings. Second, a two-stage estimation is necessary when exploring the characteristics of the latent classes. Instead of a joint maximum-likelihood analysis of the full model for the latent variable indicators, the latent continuous variables, the latent class variable, and the background variables, a second-stage analysis is needed in which estimated posterior probabilities for the latent classes are related to the background variables.

### Growth Mixture Modeling

The primary focus for Framework C in this chapter is growth mixture modeling. Verbeke and Lesaffre (1996) discussed a mixed linear model approach to random coefficient growth modeling with a mixture that allows for different means



of the random coefficients. Extending the work of Nagin and Land (1993), Nagin (1999), and Roeder, Lynch, and Nagin (1999) have discussed modeling of different trajectory classes with normal and non-normal outcomes. Nagin also recognized the importance of estimating the effects of class membership on covariates. His modeling considers fixed-effect growth within class so that all individuals are assumed to have the same within-class trajectory. His approach is termed *latent class growth mixture analysis*.

In this chapter, I consider a flexible random coefficient growth mixture model that allows within-class variation in individual trajectories. Allowing for within-class variability appears to be important in practice. Analyses of several data sets have shown that a clear-cut choice of the number of classes is not provided by BIC when requiring no within-class variability, whereas a clear-cut choice is provided when allowing within-class variability. The growth mixture model is based on Framework C. Framework C may be viewed as providing Model Framework A for each latent class category and therefore allows for considerable modeling generality. It can also be extended to Framework D, discussed next.

### Framework D: General Latent Variable Modeling

The modeling framework labeled D in Figure 10.2 is a combination of the other frameworks and is referred to as *general latent variable modeling* (LVM). This may be seen as second-generation SEM, given that the modeling capabilities are vastly enhanced relative to conventional SEM. Because the generality is gained by the introduction of both categorical and continuous latent variables, it seems appropriate to focus on the concept of latent variables and use LVM. LVM expands latent class modeling, finite mixture modeling, and SEM. On the one hand, LVM recognizes the usefulness of expanding the latent class framework. LCA concerns a particular mixture model of independence among a set of categorical latent class indicators. In contrast, LVM lets the latent classes also influence more general mixture models for the other model parts of Framework A. On the other hand, LVM recognizes the usefulness of expanding finite mixture modeling to include direct indicators of the mixture components. Finite mixture modeling infers mixture component membership from the distribution of the latent variable indicators and the background variables. In contrast, LVM can let the latent class and latent class indicator part of the model be a specific measurement model from which class membership can be inferred beyond the information from the observed variables in Framework A.

From a growth modeling perspective, the new growth modeling opportunities made available in LVM may be labeled general growth mixture modeling, as was done by B. Muthén, Brown, Khoo, Yang, & Jo (1998). In summary, GGMM goes beyond conventional random coefficient growth modeling by using

latent trajectory classes that allow for heterogeneity with respect to the influence of antecedents, growth shapes, concurrent outcomes, and later consequences. Two additional features of the analysis are particularly noteworthy: confirmatory analysis and estimation of class membership probabilities.

Confirmatory analysis is a key element in conventional SEM. A priori hypotheses can be captured by parameter restrictions. For instance, factor loadings may be fixed at zero to reflect the belief that a certain factor is not measured by a certain indicator. GGMM in Framework D offers not only the same kind of confirmatory analysis but also a second type of confirmatory analysis that concerns not hypotheses about parameters but about individuals' class membership. Confirmatory analysis with respect to parameters is very useful in growth settings. For example, there may be quadratic growth for one class, but the growth shape of a second class is hypothesized to be linear. Here, the quadratic growth factor mean and variance are fixed to zero for the second class. Or one class may have the mean of one of its growth factors constrained to be larger than that of another class. The mixture analysis benefits greatly from these types of confirmatory restrictions. In contrast, conventional finite mixture analysis is typically exploratory. Finite mixture analysis is known to sometimes give rise to numerical analysis problems, such as nonconvergence and multiple maxima of the likelihood. Confirmatory mixture analysis limits the occurrence of such problems. Recent experiences show that growth mixture analysis is a relatively well-behaved form of mixture analysis, given the multivariate trend information in the data and the possibilities for formulating confirmatory hypotheses.

Confirmatory analysis with respect to the class membership of the individuals is a feature not found in SEM. With a latent class variable, however, a researcher may want to incorporate the hypothesis that certain individuals are known to represent typical trajectories corresponding to a certain class. This knowledge may be due to auxiliary information or because the individual displays such a typical growth pattern. For example, students who are known to have been diagnosed as having reading disability in second grade may have a characteristic reading skills trajectory in the first grade. Individuals with known class membership are sometimes referred to as *training data*. As is the case of using parameter restrictions, the numerical performance of the mixture analysis benefits greatly from incorporating training data. It may be noted that multiple-group SEM corresponds to the case of all sample units contributing training data so that the latent class variable is in effect an observed categorical variable.

GGMM is a promising tool for the analysis of randomized trials. A discussion of GGMM in preventive intervention settings was given by Muthén, Brown, et al. (1998). Interventions often show an interaction between treatment and characteristics of the individual. Important individual characteristics can be captured by developmental trajectory classes. Differences between classes may refer not only to baseline characteristics, such as the initial status of a growth model (B. Muthén & Curran, 1997), but also to growth rate and, more generally,

growth shape. GGMM analysis allows the effect of treatment to vary across trajectory class and is therefore able to give a more detailed assessment of treatment effects. It is also useful to add the notion of partial compliance and use CACE estimation based on latent compliance classes. The use of GGMM in conjunction with CACE gives a flexible analysis framework for randomized trials; for some initial work see Jo and Muthén (in press).

GGMM solves a problem with piecewise growth modeling used to represent different stages of development. A weakness of piecewise growth modeling is that the point of transition from one stage to the next needs to be both known and the same for all individuals in the sample. GGMM allows more flexible piecewise growth modeling in which different unknown classes of individuals make the transition at different time points. The model would seem to be of particular interest in intervention studies in which a certain treatment may take effect after different amounts of time for different classes of individuals.

A final general comment on GGMM concerns the benefit of estimating an individual's posterior probabilities of class membership. This is the counterpart of factor scores in conventional SEM. Given an estimated model, each individual obtains a posterior probability estimate for each class computed as a function of the model parameter estimates and the individual's values on his or her observed variables. The class to which the individual most likely belongs can therefore be determined.

Estimation of class membership probabilities is of particular interest with longitudinal data. Consider as an example a model for development of reading skills through first and second grades. A GGMM analysis estimates the parameters for this development. The class membership probabilities can then be estimated for a new student before the student reaches the end of second grade using only the subset of the repeated measures available at that point in time. It is of interest to study the precision with which such probabilities can be estimated at different points in time. As an example, B. Muthén, Khoo, Francis, and Kim Boscardin (1998) conducted a study in which first-grade development was related to kindergarten precursors. This allows for a mechanism for early prediction of problematic development. The analysis approach may be particularly useful in diagnostic or preventive intervention settings. Not only may individuals who belong to different classes benefit differently from a treatment program, but also the analysis could possibly guide in the choice of treatment variation.

## Examples

### 1. Analysis of Math Achievement Development

Example 1 uses a subset of data from the Longitudinal Study of American Youth. The analyses consider females in the younger cohort measured at four time

points, Grade 7–Grade 10, beginning in 1987. Math achievement items and background information from parents on mother's education and home resources are used. Conventional random coefficient growth analysis of both girls and boys indicates that mother's education and home resources play important roles in predicting math achievement development.

In Example 1, a growth mixture model is explored for 984 girls with complete data on the variables. Figure 10.3 shows the general GGMM framework applied to this example. The latent categorical variable is represented by the circle labeled  $c$ . The intercept factor is labeled  $I$ , and the slope factor is labeled  $S$ . The analysis is first carried out without the two covariates mother's education and home resources (unconditional analysis) and then including these covariates (conditional analysis). The prediction of latent class membership from the covariates will be introduced as the last step of the analysis.

### Unconditional Analysis

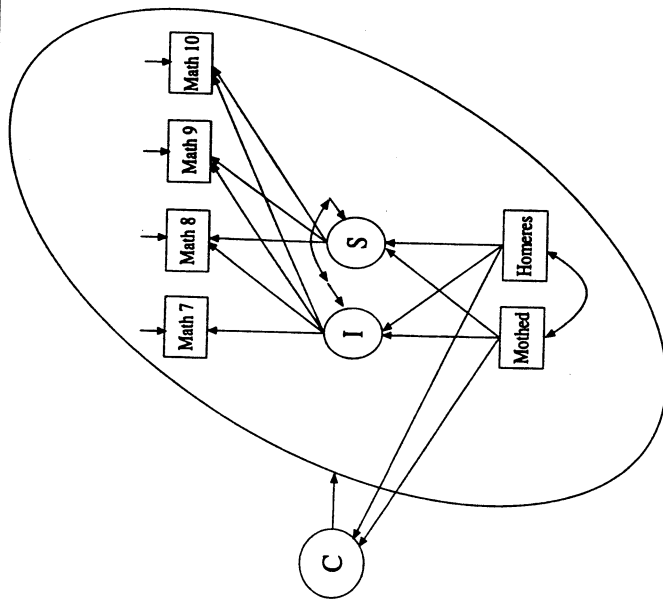
Growth mixture modeling encompasses many different model variations, and it is therefore helpful to have a good model-fitting strategy. As an initial step, one can explore the data by plotting individual observed or fitted curves. For example, with the four time points available here, a line can be estimated and plotted for each person. It may be difficult, however, to see clusters in such a set of individual curves. Two initial modeling approaches are recommended. One is to perform a conventional (one-class) growth model to get a notion of the extent of growth factor variation. Estimated individual curves can be plotted. Another approach is to search for typical classes of growth curves using latent class growth mixture analysis (LCGA), referred to in Growth Mixture Modeling. LCGA is a special case of the GGMM framework in which the growth factor covariance matrix has been fixed at zero so that there is no within-class trajectory variation. LCGA can be carried out for several different numbers of classes, producing a plot of BIC against a number of classes. The LCGA solution for the number of classes above in which no important drop in BIC occurs can be used to plot estimated mean curves. This curve plot gives an idea of how many qualitatively different numbers of classes there are; say,  $m$ . Choosing the  $m$ -class LCGA solution as a starting point, a GGMM analysis can be carried out to explore within-class variability, letting the growth factor covariance matrix be freely estimated. This model-fitting strategy will now be used. As mentioned earlier, comparison of models with different numbers of classes can be accomplished by using an information criterion such as BIC. For models that have the same number of classes and are nested, the usual likelihood-ratio chi-square difference test can be used.

The conventional one-class growth model has a BIC value of 25,534.45. The model fits well as tested against an unrestricted mean- and covariance-



FIGURE 10.3

**Math achievement modeling. Mothered = mother's education; Homeres = home resources.**

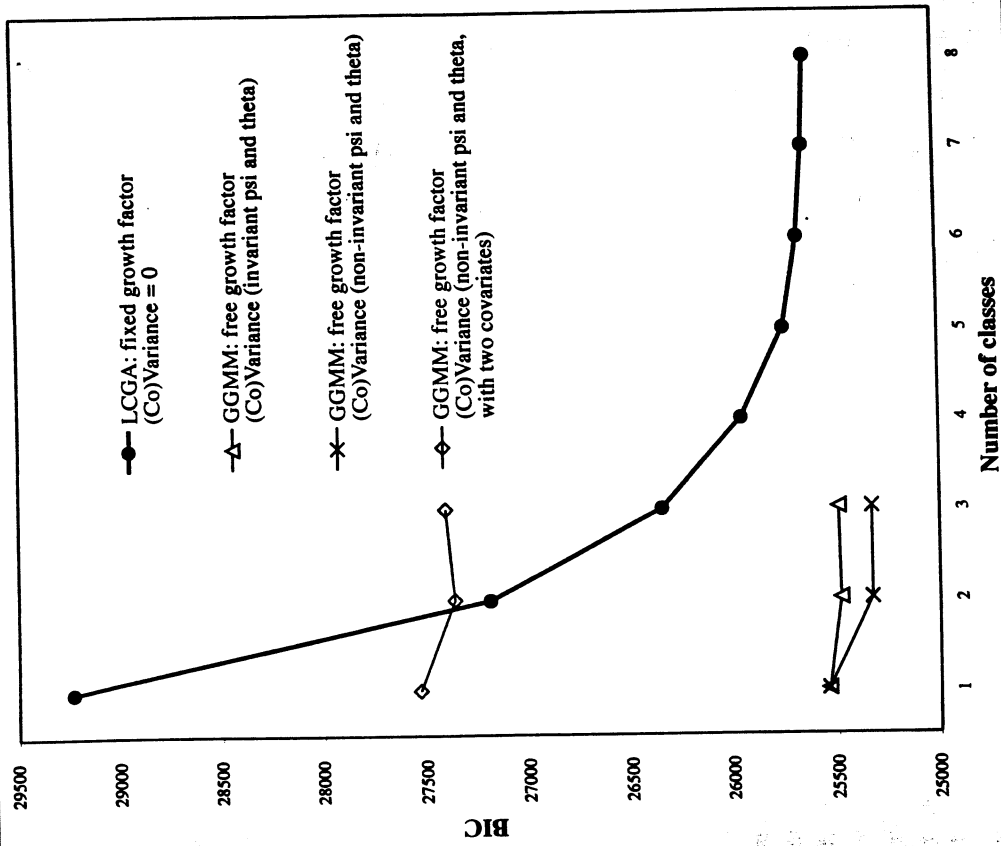


structure model:  $\chi^2(3, N = 984) = 4.22, p = .237$ , root mean square error of approximation (RMSEA) = .020 (confidence interval [CI]: .000, .061). The intercept mean, which is equal to the initial status at Grade 7, is 52.78, and the growth rate mean is 2.59. Nonlinear growth is accommodated by estimated time score values  $x_i$  at the last two grades. The growth is accelerated between Grade 8 and Grade 9, with the time score value for Grade 9 estimated as 2.45 instead of the linear growth value of 2.0. The estimated time score value for Grade 10 is 3.50. The growth factor variances are 64.50 and 1.29, respectively. Both are significantly different from zero. The growth factor correlation is .34. The residual variances for the outcomes are estimated at 14.55, 13.20, 14.13, and 26.53.

For the series of LCGA, starting values for the growth curves can be based on the conventional one-class growth model estimates, using values in the range of the growth factor means and  $\pm 2$  growth factor standard deviations. The corresponding BIC values are shown in Figure 10.4. The LCGA BIC curve levels off smoothly as the number of classes increases and does not point to a definite

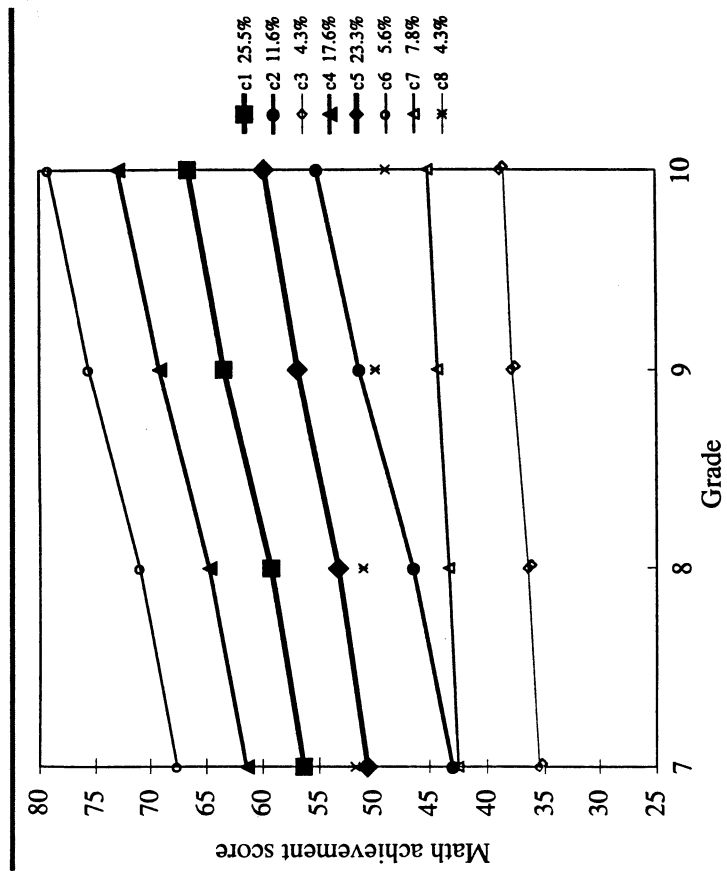
FIGURE 10.4

**Longitudinal Study of American Youth Bayesian information criteria (BIC) values for latent class growth mixture analysis (LCGA) and general growth mixture modeling (GGMM).**



number of classes as the best model. In fact, the figure shows that not even with eight classes is the LCGA BIC better than for the conventional one-class growth model. The estimated mean curves for the eight-class LCGA are shown in Figure 10.5. The more frequent the curve, the thicker the curve. Typical curves have high initial status with high growth rate and low initial status with

FIGURE 10.5  
**Longitudinal Study of American Youth mean curves for eight classes with zero factor variance.**



low growth rate. It is unclear if there are several major types of classes of curves or if there is only one class with individual variation. Figure 10.6 shows the LCGA mean curves using two and three classes. These solutions are used as starting values for GGMM analyses.

A first GGMM uses two latent classes and lets the growth factor means for the intercept factor and the slope factor vary across the classes while all other parameters are held invariant across the classes. This results in a BIC value of 25,478.22, which should be compared with the conventional single-class growth model BIC of 25,534.45 with three fewer parameters (two growth factor mean parameters and one class probability parameter). This means that the two-class model is preferable on the basis of BIC. A corresponding three-class model has a worse BIC value of 25,488.68. GGMM with two and three classes are also fitted with class-varying growth factor covariance matrix and residual variances, resulting in BIC values of 25,330.92 and 25,331.80, slightly favoring the two-class model. The BIC curves are included in Figure 10.4; see bottom

left of the figure. Figure 10.4 also includes the corresponding GGMM curve with class-varying (co)variances when adding the two covariates; see top left of the figure. The two-class model is also favored when including the covariates.

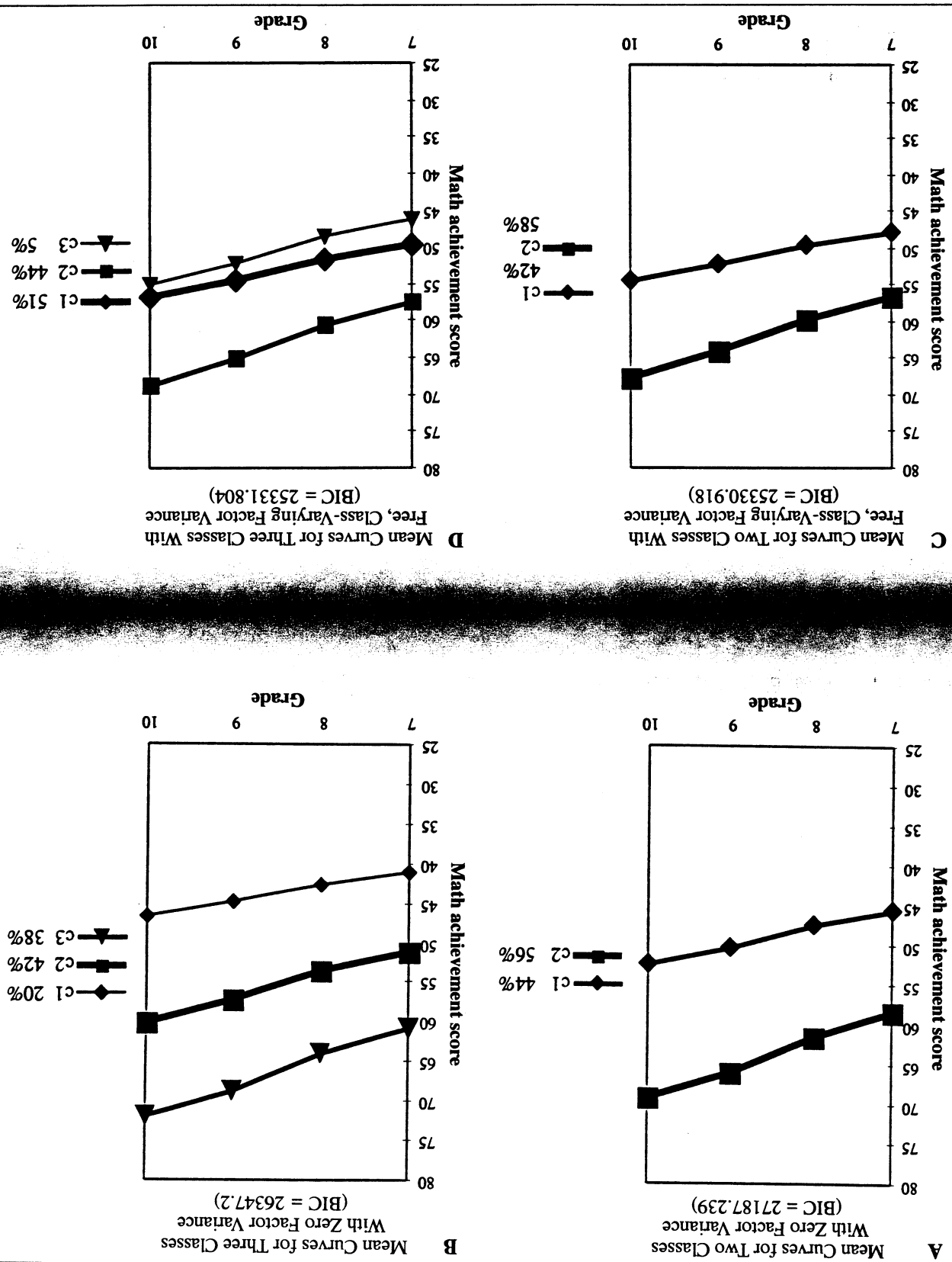
The first two-class model imposes class invariance of the covariance matrix for the two growth factors as well as class invariance of the variances for the residuals of the outcomes. Letting the variances for the residuals of the outcomes vary across the two classes gives a log likelihood value of  $-12,597.91$ . Because the two-class model with invariant residual variances has a log likelihood value of  $-12,690.87$ , the test of invariance is given by a likelihood-ratio chi-square difference test value with 4 *df* equal to 185.92. This indicates that the residual variances need to be different across classes.

Also letting the factor covariance matrix differ across the two classes gives a log likelihood value of  $-12,593.10$  and  $\chi^2(3, N = 984) = 9.62, p = .025$ , providing a weak indication that these parameters need to be different across the classes. For this model there are 42% in Class 1 which, compared with Class 2, is characterized by low starters who grow slowly. The mean curves are shown in Figure 10.6. For Class 1 the outcome residual variances are estimated as 16.42, 16.33, 22.97, and 49.06, whereas for Class 2 they are estimated as 12.52, 11.28, 8.34, and 7.80. The largest difference is for Grade 10, showing a larger degree of time-specific variance for the low-performing class. For Class 1 the intercept variance is 53.88, whereas for Class 2 it is 40.49. For Class 1 the slope variance is 1.81, whereas for Class 2 it is 0.20. All of these variances are significant except the last one mentioned. Although the growth rate variation for Class 1 is significant, the insignificant slope variance for Class 2 indicates that students in this well-performing class develop equally fast. In summary, the two classes show quite different development.

### Conditional Analysis

In the conditional analysis the growth factors are predicted using the two covariates mothers' education and home resources. The sample size is now reduced to 935. A conventional one-class growth model indicates that mother's education and home resources are both significant predictors of the intercept growth factor, whereas only home resources is significant for the slope factor. The BIC value is 27,534.74. The model fits well as judged by conventional means—the chi-square test of model fit is 15.85 with 7 degrees of freedom ( $p = .027$ ), RMSEA = .037 (CI: .012, .061). The Figure 10.4 BIC plot for the GGMM with covariates suggests a two-class model when using a class-varying factor covariance matrix and residual variances. The two-class model has a BIC value of 27,364.85. The percentage of individuals in Class 1 is estimated as 42%, the same value as for the unconditional analysis indicating a stable classification. For Class 1, mother's education is a significant predictor of the in-

Longitudinal Study of American Youth mean curves for two and three classes. BIC = Bayesian information criteria.



tercept factor but not the slope factor, whereas home resources is a significant predictor of the slope factor but not the intercept factor. For Class 2, mother's education and home resources are both significant for the intercept but not for the slope. The two-class analysis shows that the home-resource variable plays different roles in the two classes. Home resources is an important factor for the Grade 7 math achievement initial status only for students who are developing well (Class 2) and is important for the math growth rate only for those who are not developing well (Class 1).

The estimated model is given in Table 10.1. The estimated time scores are held equal across class. When allowed to vary across class, a worse BIC value is obtained: 27,375.63. The two-class results can be contrasted with the conventional one-class analysis. The conventional analysis mistakenly concludes that home resources are important for initial status and for growth for all individuals. Conventional tests of model fit indicate no misfit of the one-class model. In contrast, GGMM BIC testing suggests a mixture of two subpopulations. The two-class analysis is of educational importance in that it unmixes these two different subpopulations and carefully delineates the impact of home resources on math achievement development. An ad hoc alternative is to divide the sample into subsets of low- and high-scoring individuals in Grade 7 before the analysis. Such subsetting, however, is not recommended for three reasons. First, the subsetting would be made with respect to the observed score, which is a fallible measure of the initial status growth factor. Second, the subsetting would be made with respect to a dependent variable in the growth model, which distorts the underlying population relationships. Third, the subsetting would be made in a subjective fashion. GGMM avoids all of these problems.

As shown in the GGMM Framework C in Figure 10.3, the latent class membership may also be related to the two covariates. As part of the GGMM analysis, the logit regression of  $c$  on mother's education and home resources for the final two-class model showed no significant influence on the probability of membership in Class 1, although the logit coefficients were negative, as expected. Other, more powerful covariates may be brought in to predict class membership. Finally, posterior probabilities of class membership can be computed for each individual for further study.

## 2. Analysis of Normative and Non-Normative Development in Heavy Drinking

Example 2 uses a subset of data from the National Longitudinal Survey of Youth (NLSY). The outcome variable is frequency of heavy drinking (six or more drinks on one occasion) during the past 30 days. Covariates are gender, ethnicity, family history of alcohol problems, early start of regular drinking (age 14 or younger), dropping out of high school, and college attendance. For this

TABLE 10.1  
Estimates for Final LSAY Two-Class Model

PARAMETER	CLASS 1 (42%)		CLASS 2 (58%)	
	ESTIMATE	SE	ESTIMATE	SE
Intercept factor regression	42.73	1.65	25.92	45.37
Intercept	1.66	0.63	2.65	2.09
Motherd	0.72	0.38	1.91	1.81
Homeres	54.30	6.09	8.91	32.43
Slope factor regression	0.82	0.37	2.23	2.85
Intercept	0.15	0.15	0.95	0.03
Motherd	0.23	0.09	2.63	0.05
Homeres	1.64	0.63	2.62	0.19
Residual variance, intercept-slope	0.73	1.64	0.44	0.48
Residual variance	17.20	2.84	6.06	12.25
Grade 7	15.26	2.08	7.35	11.38
Grade 8	24.17	3.29	7.34	8.01
Grade 9	49.11	10.04	4.89	7.35
Grade 10				
Time score	2.42	0.13	0.13	18.16
Grade 9	3.50	0.20	0.20	17.57
Grade 10	-0.33	0.31	0.31	-1.06
Latent class logit				
Intercept factor regression	31.05	1.46	45.37	1.46
Motherd	1.66	0.63	2.65	2.09
Homeres	0.72	0.38	1.91	1.81
Residual variance	54.30	6.09	8.91	32.43
Slope factor regression	0.82	0.37	2.23	2.85
Intercept	0.15	0.15	0.95	0.03
Motherd	0.23	0.09	2.63	0.05
Homeres	1.64	0.63	2.62	0.19
Residual variance, intercept-slope	0.73	1.64	0.44	0.48
Residual variance	17.20	2.84	6.06	12.25
Grade 7	15.26	2.08	7.35	11.38
Grade 8	24.17	3.29	7.34	8.01
Grade 9	49.11	10.04	4.89	7.35
Grade 10				
Time score	2.42	0.13	0.13	18.16
Grade 9	3.50	0.20	0.20	17.57
Grade 10	-0.33	0.31	0.31	-1.06

Note. Bayes information criteria = 27,364.85. Number of parameters = 29. LSAY = Longitudinal Study of American Youth.

FIGURE 10.7

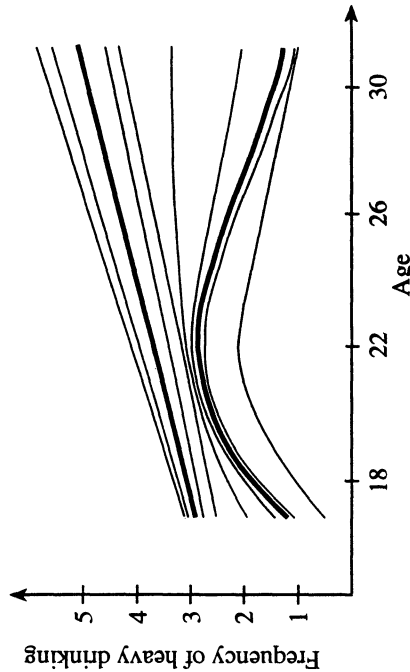
**Modeling of heavy drinking.**

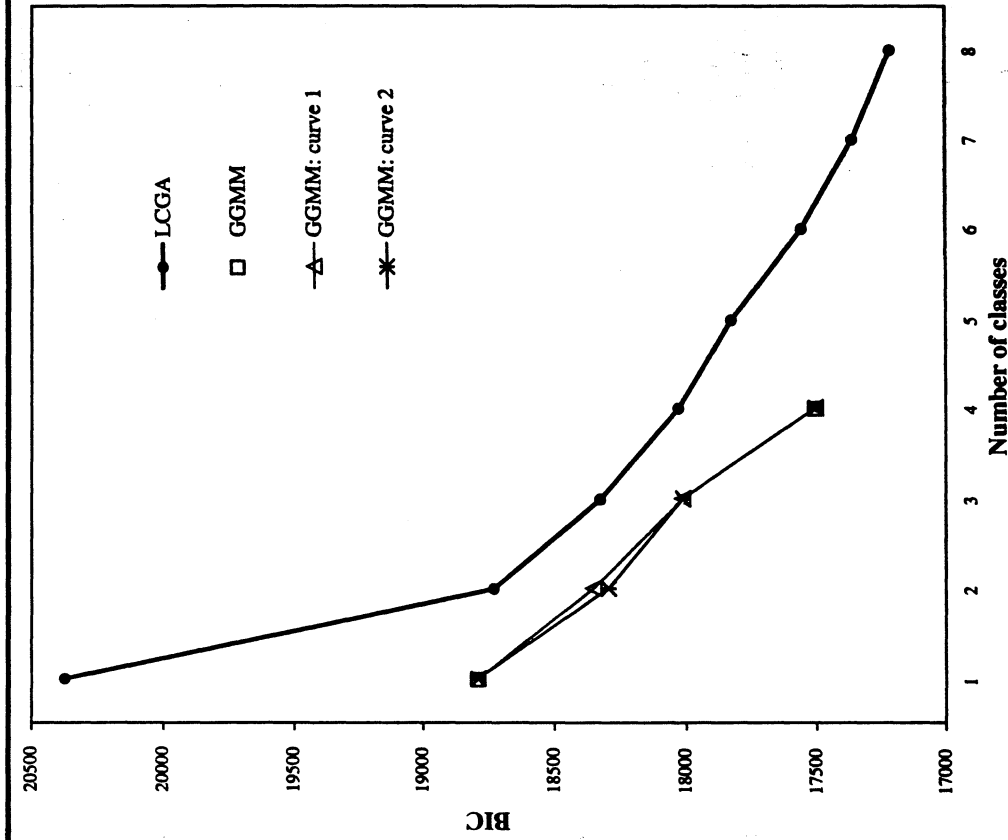
illustration heavy drinking is considered for one of eight NLSY cohorts covering ages 18, 19, 20, 24, 25, and 30. The sample size is 922.

A quadratic growth model has been found suitable for the development of heavy drinking over ages 18–37 (B. Muthén & Muthén, 2000). The growth curve shape at the average of the three growth factors is given as the bottom solid curve in Figure 10.7, showing an increase from 18 to 21 with a subsequent decrease. This normative growth curve shape has also been found for delinquent behavior and illicit drug sampling. One may ask if there are also other, non-normative, growth curve shapes represented in this population. For example, some individuals may show an increase from 18 to 21 but no downturn as illustrated by the solid line. Given that the conventional, single-class, quadratic random coefficient model allows for individual variation in all three growth factors, the solid line trajectory shape is actually included as a special case in the conventional model. The question is whether a better fit to the data can be obtained by a two-class growth model in which the classes differ in their means on the three growth factors. In line with the GGMM model Framework C, such a two-class model can be used to estimate the influence of covariates on the probability of class membership. The model can also be used to estimate each individual's most likely trajectory class membership.

**Growth Curve Shapes**

B. Muthén and Shedden (1999) performed GGMM analyses on the NLSY data.<sup>14</sup> The same model search strategy as in Example 1 can be carried out. The BIC plot for LCGA with 1–8 classes is shown in Figure 10.8 together with BIC

FIGURE 10.8

**National Longitudinal Survey of Youth Bayesian information criteria (BIC) values for latent class growth mixture analysis (LCGA) and general growth mixture modeling (GGMM).**


values for GGMM with 1–4 classes. The two-class GGMM resulted in two different solutions; the two solutions shared the normative curve for the majority but had different non-normative curves. A three-class solution showed all three kinds of curve shapes and was also found to fit the data better in terms of BIC value. A three-class solution in line with Muthén and Shedden is shown

in Figure 10.9. The normative class probability was estimated as 77%, whereas the two non-normative class probabilities were estimated as 14% and 9%, respectively. The four-class solution has an even better BIC value, with two classes starting high at age 18, but does not alter the general conclusions and is not presented here.

As part of the GGMM analysis, membership in the three classes is also related to the set of covariates. Table 10.2 shows the estimates of the corresponding multinomial logistic regression. In the table, *High* refers to the class in Figure 10.9 that is high already at age 18, *Increase* refers to the class that increases over age, and *Norm* refers to the normative class. In the column labeled *High vs. norm* the coefficients show how the odds of belonging to the high class compared to the normative class is significantly increased for individuals who are male and have early onset and is significantly decreased for individuals who are Black and did not go to college. In the column labeled *Increase vs. norm*, the coefficients show how the odds of belonging to the increasing class compared to the normative class is significantly increased for individuals who are male, have a family history of alcohol problems, and dropped out of high school.

### Growth Curve Shapes as Predictors of Distal Outcomes

The three-class model is now embedded in a larger model shown in Figure 10.10. The model will now be used to predict a binary variable of alcohol dependence at age 30. The latent trajectory class variable in Figure 10.10 is seen to influence the growth curve shape of heavy drinking development by means of the growth factor means, and the latent trajectory class variable is predicted by covariates. The additional feature in Figure 10.10 is the prediction of alcohol dependence by trajectory class. This is a logistic regression relationship, although the predictor variable is latent. In Figure 10.2 terms, alcohol dependence is an example of a latent class indicator variable, and the model is an example of Model Framework D.

The use of a latent trajectory class variable as a predictor solves a dilemma that occurs when one tries to use the continuous growth factors as predictors. Consider as an example the two trajectory classes of Figure 10.7. Here the growth factor representing linear growth rate concerns the rate of increase right after age 18. This linear growth rate factor is not, however, a suitable predictor of later problematic alcohol use outcomes, because both classes have a high positive slope value. Likewise, if there is a third class that is high already at age 18, a nonpositive slope may be predictive of problematic outcomes. The key issue is that the growth factor values interact in determining the growth shape, and it is the shape that is predictive.

Table 10.3 shows the resulting estimates of the relation between trajectory

FIGURE 10.9  
Estimated heavy-drinking curves.

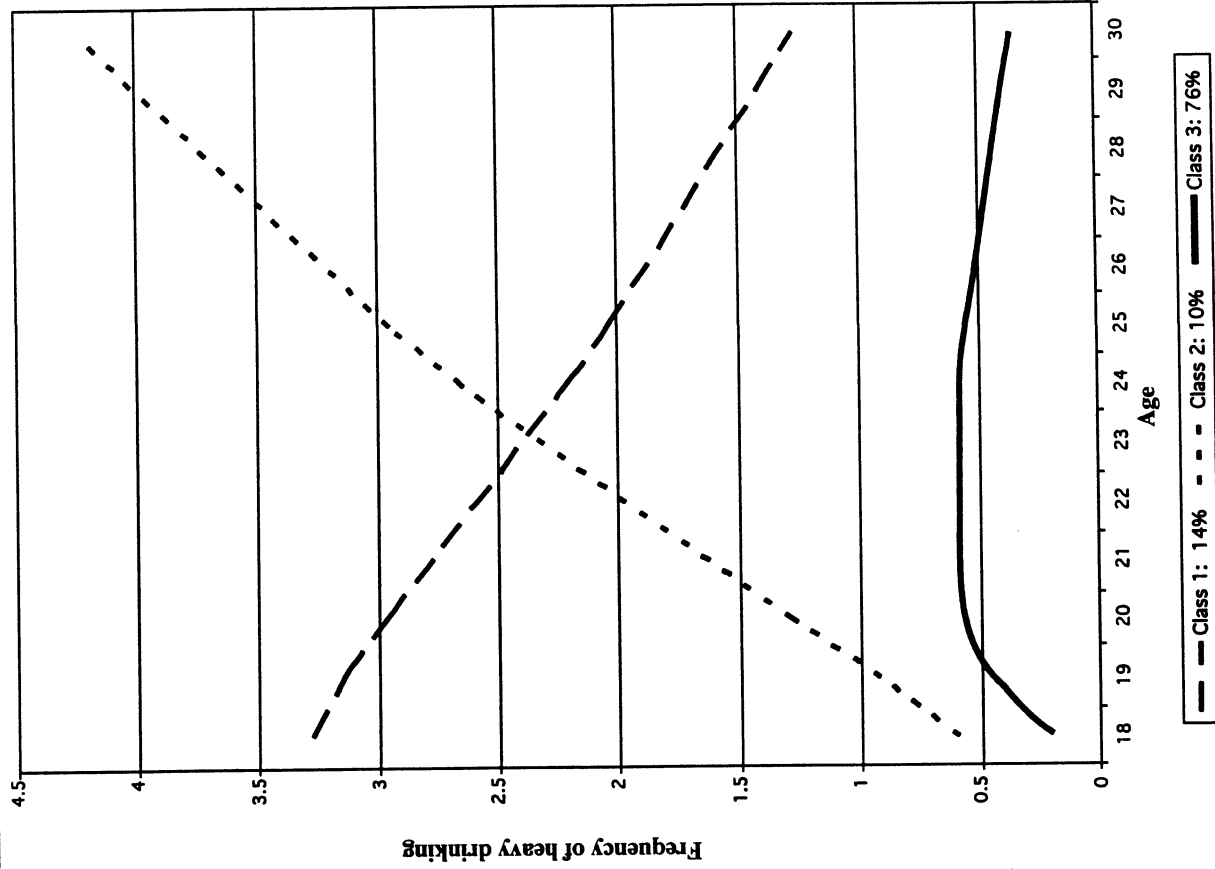




TABLE 10.2  
**Predicting Trajectory Class Membership in the NLSY:  
 Estimated Logit Coefficients**

COVARIATE	HIGH VS. NORM	INCREASE VS. NORM
Male	1.45*	1.50*
Black	-1.94*	-0.14
Hispanic	-0.49	-0.10
Early onset	1.49*	0.78*
FH123	0.45	1.07*
Dropout	0.37	0.77*
College	-0.63*	-0.92*

Note. NLSY = National Longitudinal Survey of Youth; FH123 = Family history of alcohol problems among first- and second- or third-degree relatives.  
 \* $p < .05$ .

class and alcohol dependence. Although the normative class ( $c = 3$ ) has an estimated probability of .05 of developing alcohol dependence by age 30, the two non-normative classes have elevated probabilities of .59 ( $c = 2$ ) and .20 ( $c = 1$ ), respectively. Table 10.3 also gives the odds of developing alcohol dependence in each class and the corresponding odds ratios when comparing a class with

FIGURE 10.10

**Modeling of heavy drinking: Predicting a distal outcome. FH = family history of alcohol problems; ASB = antisocial behavior; HSDrp = high school dropout.**

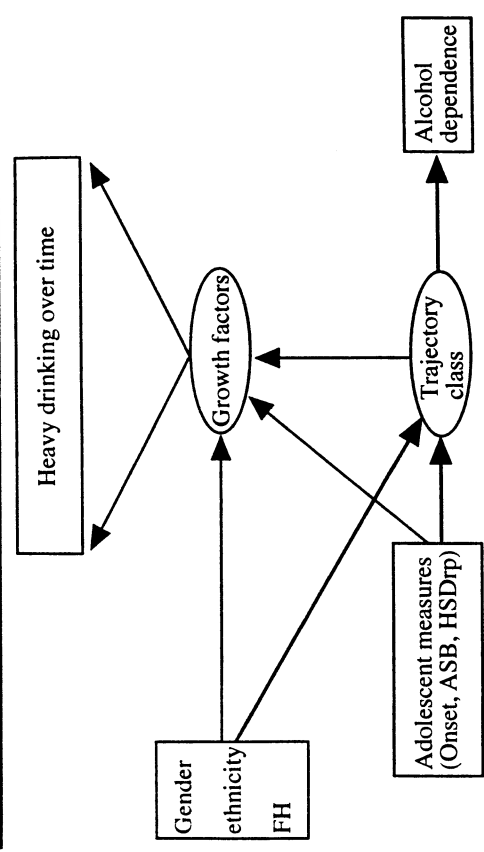


TABLE 10.3

**Probability of Dependence Given Class**

PROBABILITY FUNCTION	ODDS	ODDS RATIO
$P(\text{Dep}   c = 1) = .20$	0.26	5.19
$P(\text{Dep}   c = 2) = .59$	1.43	29.07
$P(\text{Dep}   c = 3) = .05$	0.05	1.00

the normative class. Given that the probabilities vary so strongly as a function of the latent trajectory class membership, it is of great importance to be able to predict the class membership as early as possible using both covariate information and early information on heavy-drinking development.

**Conclusion**

It is clear from the examples that GGMM is an important new development for the study of change. The new modeling opportunities will enrich growth modeling and allow more complex ideas of development based on substantive theories in various fields. GGMM is strengthened by the fact that it is part of a general latent variable framework. In this way, it is likely that growth modeling and latent-variable modeling will each continue to benefit from developments in the other's area.

**References**

Agresti, A. (1990). *Categorical data analysis*. New York: Wiley.

Arminger, G., & Stein, P. (1997). Finite mixtures of covariance structure models with regressors. *Sociological Methods and Research*, 26, 148-182.

Arminger, G., Stein, P., & Wittenberg, J. (1999). Mixtures of conditional mean- and covariance-structure models. *Psychometrika*, 64, 475-494.

Bandeau-Roche, K., Miglioretti, D. L., Zeger, S. L., & Rathouz, P. J. (1997). Latent variable regression for multiple discrete outcomes. *Journal of the American Statistical Association*, 92, 1375-1386.

Bartholomew, D. J. (1987). *Latent variable models and factor analysis*. New York: Oxford University Press.

Bläfield, E. (1980). *Clustering of observations from finite mixtures with structural information*. Unpublished doctoral dissertation, Jyväskylä Studies in Computer Science, Economics, and Statistics, Jyväskylä, Finland.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

- 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.
- Collins, L. M., & Wugalter, S. E. (1992). Latent class models for stage-sequential dynamic latent variables. *Multivariate Behavioral Research, 27*, 131–157.
- Dayton, C. M., & Macready, G. B. (1988). Concomitant variable latent class models. *Journal of the American Statistical Association, 83*, 173–178.
- Everitt, B. S., & Hand, D. J. (1981). *Finite mixture distributions*. London: Chapman & Hall.
- Formann, A. K. (1992). Linear logistic latent class analysis for polytomous data. *Journal of the American Statistical Association, 87*, 476–486.
- Gibson, W. A. (1959). Three multivariate models: Factor analysis, latent structure analysis, and latent profile analysis. *Psychometrika, 24*, 229–252.
- 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.
- Jedidi, K., Ramaswamy, V., DeSarbo, W. S., & Wedel, M. (1996). On estimating finite mixtures of multivariate regression and simultaneous equation models. *Structural Equation Modeling, 3*, 266–289.
- Jo, B., & Muthén, B. (in press). Intervention studies with noncompliance: Complicating causal effect estimation in growth mixture modeling. In N. Duan & S. Reise (Eds.), *Multilevel modeling: Methodological advances, issues, and applications*. Mahwah, NJ: Erlbaum.
- Little, R. J., & Yau, L. H. Y. (1998). Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin's causal model. *Psychological Methods, 3*, 147–159.
- McLachlan, G. J., & Basford, K. E. (1988). *Mixture models: Inference and applications to clustering*. New York: Marcel Dekker.
- Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika, 55*, 107–122.
- Moffitt, T. E. (1993). Adolescence-limited and life-course persistent antisocial behavior. *Psychological Review, 100*, 674–701.
- Muthén, B., Brown, C. H., Khoo, S., Yang, C., & Jo, B. (1998, June). *General growth mixture modeling of latent trajectory classes: Perspectives and prospects*. Paper presented at the annual meeting of the Prevention Science and Methodology Group, Tempe, AZ.
- Muthén, B., & Curran, P. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 2*, 371–402.
- Muthén, B., Khoo, S. T., Francis, D., & Kim Boscardin, C. (1998). *Analysis of reading skills development from kindergarten through first grade: An application of growth mixture modeling to sequential processes*. Unpublished manuscript, University of California, Los Angeles.
- Muthén, B., & Muthén, L. (2000). The development of heavy drinking and alcohol-related problems from ages 18 to 37 in a U.S. national sample. *Journal of Studies on Alcohol, 61*, 290–300.
- Muthén, B., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics, 55*, 463–469.
- Muthén, B., Shedden, K., & Spisic, D. (1998). *General latent variable mixture modeling*. Unpublished manuscript, University of California, Los Angeles.
- Muthén, L. K., & Muthén, B. (1998). *Mplus user's guide*. Los Angeles, CA: Muthén & Muthén.
- Nagin, D. S. (1999). Analyzing developmental trajectories: A semi-parametric, group-based approach. *Psychological Methods, 4*, 139–157.
- Nagin, D., Farrington, D., & Moffitt, T. (1995). Life-course trajectories of different types of offenders. *Criminology, 33*, 111–139.
- Nagin, D. S., & Land, K. C. (1993). Age, criminal careers, and population heterogeneity: Specification and estimation of a nonparametric, mixed Poisson model. *Criminology, 31*, 327–362.
- Pearson, J. D., Morrell, C. H., Landis, P. K., Carter, H. B., & Brant, L. J. (1994). Mixed-effect regression models for studying the natural history of prostate disease. *Statistics in Medicine, 13*, 587–601.
- Roeder, K., Lynch, K. G., & Nagin, D. S. (1999). Modeling uncertainty in latent class membership: A case study in criminology. *Journal of the American Statistical Association, 94*, 766–776.
- Schulenberg, J., O'Malley, P. M., Bachman, J. G., Wadsworth, K. N., & Johnston, L. D. (1996). Getting drunk and growing up: Trajectories of frequent binge drinking during the transition to young adulthood. *Journal of Studies on Alcohol, 57*, 289–304.
- Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.
- Titterton, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions*. Chichester, UK: Wiley.
- van der Heijden, P. G. M., Dessens, J., & Böckenholt, U. (1996). Estimating the concomitant-variable latent-class model with the EM algorithm. *Journal of Educational and Behavioral Statistics, 21*, 215–229.
- Verbeke, G., & Lesaffre, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population. *Journal of the American Statistical Association, 91*, 217–221.
- Yung, Y. F. (1997). Finite mixtures in confirmatory factor-analysis models. *Psychometrika, 62*, 297–330.

Zucker, R. A. (1994). Pathways to alcohol problems and alcoholism: A developmental account of the evidence for multiple alcoholisms and for contextual contributions to risk. In R. A. Zucker, J. Howard, & G. M. Boyd (Eds.), *The development of alcohol problems: Exploring the biopsychosocial matrix of risk* (NIAAA Research Monograph No. 26, pp. 255-289). Rockville, MD: U.S. Department of Health and Human Services.