Multilevel Modeling

Methodological Advances, Issues, and Applications

Edited by

Steven P. Reise
Naihua Duan
University of California, Los Angeles

2003
LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London

6

Longitudinal Studies With Intervention and Noncompliance: Estimation of Causal Effects in Growth Mixture Modeling

Booil Jo
Bengt O. Muthén
UCLA

The major interest in intervention trials is often the estimation of intervention effects for individuals who actually receive the intervention. However, some percentage of noncompliance is usually unavoidable in intervention trials when dealing with human participants. In addition, it is not easy to control compliance behavior of individuals who may decide not to participate even with highly attractive incentives. Noncompliance is a major threat to obtaining power to detect intervention effects (Jo, 2002), and may bias the estimation of intervention effects if not handled carefully in the statistical analysis.

ITT (intent to treat) analysis is a standard way to estimate intervention effects in randomized experimental designs in the presence of noncompliance. In this method, average outcomes are compared by randomized groups, ignoring the existence of noncompliance. Because the standard ITT analysis often underestimates intervention effects in the presence of noncompliance, the possibility of estimating intervention effects only for the individuals who actually receive the intervention has been explored using CACE (complier average causal effect) estimation (Angrist, Imbens, & Rubin, 1996; Bloom, 1984; Imbens & Rubin, 1997; Little & Yau, 1998). In CACE approaches, compliers and noncompliers are allowed to be different in various aspects and are best thought of as belonging to different subpopulations. For example, people with higher motivation or a special interest in the intervention will be more likely to comply with the intervention.

In the CACE estimation method, the causal effect of intervention is usually defined based on a single outcome observed after intervention, treating the baseline measure as one of the covariates (i.e., ANCOVA approach). However, when an intervention study is focused on the long-term effect of the intervention, the outcome is often measured several times at specific intervals. In this case, considering the longitudinal nature of intervention studies, it is also possible, and perhaps more natural, to define the intervention effect based on a trend or a growth trajectory of individuals. This study demonstrates CACE estimation based on latent trajectories over time in a growth mixture modeling framework.

One advantage of using a growth modeling framework is that the first time point measure is considered as one of the outcome measures instead of as one of the covariates. This parameterization adds more flexibility in the interpretation of the results because initial status and growth rate of outcome measures are separated. For example, the influence of background variables can be estimated separately for initial status and growth rate of the outcome measure.

Another advantage of this model is that it utilizes not only covariate but also trajectory information, which often improves precision in estimating the compliance type of individuals. Including growth process in the estimation of CACE utilizes the idea of a general latent variable modeling framework where both categorical and continuous latent variables are incorporated (Muthén, 2001a; Muthén, 2001b, Muthén & Muthén, 1998-2001). That is, latent variables that represent growth trajectories are continuous as in conventional structural equation models, whereas the latent variable that represents compliance status is categorical. To differentiate growth modeling with both categorical and continuous latent variables from traditional growth modeling, the former will be called "growth mixture modeling" in this study. This study focuses specifically on random coefficient growth mixture modeling where individual variation is allowed within each class or compliance status. In contrast, individual variation is not allowed within each growth trajectory class in a group-based modeling approach (Nagin, 1999).

This study also explores the possibility of using exploratory growth mixture analysis as a data mining tool that precedes growth mixture CACE analysis. Growth mixture CACE analysis is considered confirmatory because compliance type is known for individuals who are assigned to the intervention condition. In exploratory growth mixture analysis, individuals are classified into several classes without observed class information (training data), and efficiency of classification can often be improved by utilizing the fact that certain trends are present in longitudinal data (Muthén, Brown, Khoo, Yang, & Jo, 1997; Muthén et al., in press; Muthén & Shedden, 1999). When the intervention condition includes many sessions, or doses, one needs to determine the appropriate cutpoints that separate individuals into different classes based on level of compliance. Exploratory growth mixture analysis can be useful in determining cutpoints at the planning stage of growth mixture CACE analysis.

This chapter is organized as follows. Section 1 describes the estimation method using the ML-EM algorithm and defines model assumptions in the estimation of CACE in this study. Section 2 demonstrates the efficiency of CACE estimation in growth mixture modeling through simulation studies. Section 3 demonstrates how exploratory and confirmatory growth mixture analyses can be used in studying unknown subpopulations using the Johns Hopkins Preventive Intervention Study in Baltimore Public Schools as an example. Section 4 concludes with discussion.

ESTIMATING DIFFERENTIAL EFFECTS OF INTERVENTIONS

Model assumptions

The CACE models used in this study are based on statistical assumptions in line with Rubin's causal model. In Rubin's causal model approach, the possibility of statistical causal inference is built based on the effect of treatment assignment at the individual level (Holland, 1986; Rubin, 1974, 1978, 1980). Stable unit treatment value (SUTVA) implies that potential outcomes for each person are unrelated to the treatment status of other individuals (Rubin, 1978, 1980, 1990).

SUTVA and randomization in the study provide a statistical means of causal inference at the population level. Based on these assumptions, four types of subpopulations can be defined by classifying the potential behavior types of the subjects. Angrist et al. (1996) labeled the four categories as complier, never-taker, defier, and always-taker based on assignment and receipt of treatment. Compliers are subjects who do what they are assigned to do. Never-takers are subjects who do not receive the treatment even if they are assigned to the treatment condition. Defiers are the subjects who do the opposite of what they are assigned to do. Always-takers are the subjects who always receive the treatment no matter which condition they are assigned. Among these four kinds of possible compliance types, the current study assumes only two types of compliance and eliminates the possibility of defiers and always-takers. The assumption of monotonicity (Imbens & Angrist, 1994) excludes the possibility of having defiers. In addition, the current study assumes that there are no always-takers, which is the case when study participants are prohibited from receiving a different intervention condition than the one to which they were assigned as in the real data examples shown in later sections. However, unlike monotonicity, the assumption of having no always-takers is not critical in estimating the CACE and can be relaxed depending on the situation.

Unlike ITT analysis, CACE analysis involves methodological complexities due to the missingness of compliance information among control condition individuals. In conventional CACE approaches, it is assumed that the outcome is independent of the treatment assignment for never-takers and always-takers (the exclusion restriction assumption, Angrist et al., 1996). This assumption plays a critical role in simplifying methodological difficulties involved in CACE approaches. Under this assumption, treatment effects are estimated for compliers, but are fixed at zero for the rest. However, this assumption can be unrealistic in some situations (Hirano, Imbens, Rubin, & Zhou, 2000; Jo, in press-a, b). In the Johns Hopkins Preventive Intervention Study example shown in this study, it seems more reasonable to dichotomize individuals as low compliers and high compliers than as never-takers and compliers. In this case, it is possible that the intervention might have a weaker impact on low compliers, but it could not be guaranteed that the intervention has no effect at all, because low compliers were also exposed to the intervention.

Growth mixture CACE modeling using ML-EM estimation method

This study focuses on average causal effect estimation in the random coefficient growth mixture modeling framework. In this study, CACE estimation is used to refer to a more general method that differentiates average causal effect at varying levels of compliance, although the CACE method usually means causal effect estimation that is limited only to compliers.

In growth mixture analysis, the observed outcome variable can be expressed in terms of continuous latent variables that capture growth trajectories over time. Consider a single outcome variable y for individual

i at time point t,

$$y_{it} = I_{ik}\,\lambda_{It} + S_{ik}\,\lambda_{St} + \epsilon_{it}, \tag{6.1}$$

where latent categorical variable $\mathbf{c}$ has K levels of compliance status ($k = 1, 2, ..., K$). Compliance status $\mathbf{c}$ is observed in the intervention group and latent (missing) in the control group. Variable $c_i = (c_{i1}, c_{i2}, ..., c_{ik})$ has a multinomial distribution, where $\mathbf{c}_{ik} = 1$ if individual i belongs to class k and zero otherwise. The categorical latent variable approach may also be referred to as finite mixture modeling where sampling units consist of subpopulations that might have separate distributions and different model parameters (McLachlan & Peel, 2000; Titterington, Smith, & Makov, 1985). In finite mixture modeling, the number of mixture components is assumed to be known and fixed. For example, $K = 2$ in simulation studies and real data examples shown in later sections. Here, I_{ik} and S_{ik} are individually varying continuous latent variables representing initial level of outcome and growth rate (slope) respectively. The time scores for the initial status (λ_{It}) are equal across all time points (usually fixed at 1.0) because initial status does not change over time. The time scores for the growth rate (λ_{St}) are 0, 1, 2,..., T, representing linear growth over time, which may be fixed at different values depending on the distance between the measuring points. And ϵ_{it} represents a normally distributed residual at time point t with zero mean and variance ${\sigma_t}^2$.

Individual variation in growth parameters I_{ik} and S_{ik} within compliance class k can be expressed as

$$I_{ik} = I_k + \gamma_{Ix}\,\mathbf{x}_i + \zeta_{Iik}, \tag{6.2}$$
$$S_{ik} = S_k + \gamma_{Sx}\,\mathbf{x}_i + \gamma_{Zk}\,Z_i + \zeta_{Sik}, \tag{6.3}$$

where I_k and S_k represent intercept parameters of initial status and slope for each compliance class k; $\mathbf{x}$ represents a vector of observed covariates, and γ_{Ix} and γ_{Sx} are regression coefficient parameters. And ζ_{Iik} and ζ_{Sik} are normally distributed residuals with zero means and variances ψ_{Ik}, ψ_{Sk}, and a covariance ψ_{ISk}. The Z_i is a binary variable that represents intervention assignment, where $Z_i = 1$ if individual i is assigned to the intervention condition and zero if individual i is assigned to the control condition. Based on randomization, growth rate (slope) is regressed on Z_i, but initial status is not regressed on Z_i. The γ_{Zk} represents a mean shift in the slope when subject i belongs to the intervention condition, and is allowed to vary across different compliance status. In this study, intervention effect is defined as the difference between intervention and control conditions in the outcome measure at the final time point. Based on Equations 6.1, 6.2 and 6.3, the average causal effect (ACE) of an intervention assignment can be defined at compliance level k at the last time point (i.e., $\lambda_{St} = T$) as

$$ACE_k = \gamma_{Zk} \times T. \tag{6.4}$$

The class probability $\boldsymbol{\pi}_i$ is allowed to vary as a function of covariates. When background variables are available, the multinomial logit model of $\boldsymbol{\pi}_i$ with a vector of covariates $\mathbf{x}$ is decribed as

$$logit(\boldsymbol{\pi}_i) = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \, \mathbf{x}_i, \tag{6.5}$$

where $\boldsymbol{\pi}_i$ is a K dimensional vector $(\pi_{i1}, \pi_{i2}, ..., \pi_{iK})'$, $\pi_{ik} = P(c_{ik} = 1 \mid \mathbf{x}_i)$, and $logit(\boldsymbol{\pi}_i) = [log(\pi_{i1}/\pi_{iK}), log(\pi_{i2}/\pi_{iK}), ..., log(\pi_{i,K-1}/\pi_{iK})]'$. The $\boldsymbol{\beta}_0$ are $K-1$ dimensional logit intercepts, and $\boldsymbol{\beta}_1$ are multinomial logit regression coefficient parameters. The multinomial logit regression also provides information about the characteristics of individuals with different compliance levels.

CACE analyses reported in this study were carried out by the M*plus* program (Muthén & Muthén, 1998-2001) using maximum likelihood estimation via the EM algorithm (Dempster, Laird, & Rubin, 1977; Little & Rubin, 1987; McLachlan & Krishnan, 1997; Tanner, 1996). In the ML-EM method, the unknown compliance status (c) is handled as missing data.

Consider the sampling distribution of $\mathbf{y}$ and $\mathbf{x}$ from the mixture of k components

$$g(\mathbf{y}, \mathbf{x} \mid \boldsymbol{\theta}, \boldsymbol{\pi}) = \sum_{k=1}^{K} \pi_k f(\mathbf{y}, \mathbf{x} \mid \boldsymbol{\theta}_k), \tag{6.6}$$

where $\mathbf{y}$ and $\mathbf{x}$ represent observed data, $\boldsymbol{\theta}$ represents model parameters, and π_k represents the proportion of the population from component k with $\sum_{k=1}^{K} \pi_k = 1$. The probability $\boldsymbol{\pi}$ is the parameter that determines the distribution of $\mathbf{c}$. The observed-data log likelihood is

$$LogL = \sum_{i=1}^{n} log(\mathbf{y}_i \mid \mathbf{x}_i). \tag{6.7}$$

Given the formulation of the proposed growth mixture CACE model, the complete-data log likelihood can be written as

$$LogL_c = \sum_{i=1}^{n} [log(\mathbf{c}_i \mid \mathbf{x}_i) + log(\boldsymbol{\eta}_i \mid \mathbf{c}_i, \mathbf{x}_i) + log(\mathbf{y}_i \mid \boldsymbol{\eta}_i)], \tag{6.8}$$

where

$$\sum_{i=1}^{n} log(\mathbf{c}_i \mid \mathbf{x}_i) = \sum_{i=1}^{n} \sum_{k=1}^{K} c_{ik} \, log \, \pi_{ik}. \tag{6.9}$$

In Equations 6.8 and 6.9, $\mathbf{c}$ represents categorical latent compliance class, and $\boldsymbol{\eta}$ represents continuous latent growth factors (i.e., I and S).

Maximum likelihood estimation using the EM algorithm considers complete-data log likelihood shown in Equation 6.8. The E step computes the expected values of the complete data–sufficient statistics, given data and current parameter estimates. Compliance status $\mathbf{c}$ is considered as missing data in this step. The conditional distribution of $\mathbf{c}$, given the observed data and the current value of model parameter estimates $\boldsymbol{\theta}$, is given by

$$f(\mathbf{c} \mid \mathbf{y}, \mathbf{x}, \boldsymbol{\theta}) = \prod_{i=1}^{n} f(\mathbf{c}_i \mid \mathbf{y}_i, \mathbf{x}_i, \boldsymbol{\theta}). \tag{6.10}$$

The E step applies to both confirmatory (i.e., CACE) and exploratory growth mixture analyses, but the difference is that growth mixture CACE analysis uses information about already–known class membership (i.e., compliance status) in the intervention condition. Therefore, the first step of the implentation is easily modified with a known value of the indicator c_{ik}.

The M step computes the complete data ML estimates with complete data–sufficient statistics replaced by their estimates from the E step. This procedure continues until it reaches optimal status. The M step maximizes

$$\sum_{i=1}^{n} \sum_{k=1}^{K} p_{ik} \, log \, \pi_{ik} \tag{6.11}$$

with respect to model parameters. The p_{ik} is the posterior class probability of individual i, conditioning on observed data and model parameters, where $\pi_{ik} = P(c_{ik} \mid \mathbf{x}_i)$.

The identifiability and precision of mixture and growth mixture models used for CACE analyses in this study are based on observed compliance class membership in the intervention condition (training data) and various sources of auxilliary information such as from covariates and growth trajectories. For more details about identifiability and efficiency of extended CACE models, see Jo (in press-a). Parametric standard errors are computed from the information matrix of the ML estimator using both the first– and the second–order derivatives under the assumption of normally distributed outcomes. For more details about estimation procedures in growth mixture modeling, see Muthén & Muthén (1998-2001), Muthén & Shedden (1999), and Muthén et al. (in press).

CACE ESTIMATION USING THE OBSERVED AND LATENT VARIABLES: SIMULATION STUDIES

In longitudinal intervention studies, the effect of the intervention can be defined as the difference between the intervention and the control group in the observed outcome measured at the last time point, conditioning on the outcome measured at the first time point (ANCOVA approach). An alternative is to define the intervention effect based on latent variables that capture growth trajectories of individuals (growth model approach) as described in the previous section. This section demonstrates the quality of average causal effect estimates based on observed variable (ANCOVA) and latent variable (growth model) approaches. The simulation studies shown in this section assume that there are two underlying subpopulations with different compliance behaviors ($K = 2$). One subpopulation consists of individuals who would show a high level of compliance if assigned to the intervention condition (high compliers). The other subpopulation consists of individuals who would show a low level of compliance if assigned to the intervention condition (low compliers). The ratio of high and low compliers is 50:50, and the ratio of individuals assigned to the intervention and control conditions is 50:50. It is assumed that the intervention assignment is binary (intervention condition if $Z = 1$, control condition if $Z = 0$) and has differential effects on high compliers and low compliers. The true parameter values, effect size, and sample size are chosen based on the Johns Hopkins Public School Preventive Intervention Study example that is shown in a later section. Covariates are not included in this setting.

The true initial status mean (I_h) and the true mean growth rate (S_h) for high compliers are

$$\begin{pmatrix} I_h \\ S_h \end{pmatrix} = \begin{pmatrix} 5.00 \\ -0.25 \end{pmatrix}.$$

The true initial status mean (I_l) and the true mean growth rate (S_l) for low compliers are

$$\begin{pmatrix} I_l \\ S_l \end{pmatrix} = \begin{pmatrix} 3.50 \\ 0.15 \end{pmatrix}.$$

The true additional growth rates for high compliers (γ_{Zh}) and low compliers (γ_{Zl}) when they are assigned to the intervention condition are $\gamma_{Zh} = 0.20$, $\gamma_{Zl} = -0.10$, where the positive value of γ_{Zh} represents a desirable effect of the intervention for high compliers, and the negative value of γ_{Zl} represents a negative effect of the intervention for low compliers, assuming that positive growth of the outcome is desirable.

The true initial status variance is the same for high compliers and low compliers ($\psi_{Ih} = \psi_{Il} = 0.64$), and the true growth rate residual variance is the same for high compliers and low compliers ($\psi_{Sh} = \psi_{Sl} = 0.0625$). The true residual covariance between initial status and growth rate is zero for both high compliers and low compliers ($\psi_{ISh} = \psi_{ISl} = 0$), but is not fixed at zero in the analyses. Both variances and covariances are assumed to be equal across high and low compliers in the analyses. However, in real data examples shown in a later section, both variances and covariances are allowed to vary across high and low compliers.

It is assumed in the simulation setting that the outcome is measured four times with equal distances and has a linear trend over time. The initial status does not change over time. Given that, the fixed time scores for initial status and growth rate used in both data generation and growth mixture CACE analyses are

$$\left(\ \lambda_{It},\ \ \lambda_{St}\ \right) = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}.$$

The true residual variances and covariances of observed outcome measures are

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{pmatrix} = \begin{pmatrix} 0.36 & 0 & 0 & 0 \\ 0 & 0.49 & 0 & 0 \\ 0 & 0 & 0.64 & 0 \\ 0 & 0 & 0 & 1.00 \end{pmatrix},$$

where true residual covariances are zero, and are also assumed to be zero in the analyses. The true residual variances of outcome measures result in R^2 of 0.64 at the first time point, and 0.55 at the last time point.

The model used for data generation and the CACE analysis using the latent variable approach can be described as

$$y_{it} = I_{ik}\,\lambda_{It} + S_{ik}\,\lambda_{St} + \epsilon_{it}, \tag{6.12}$$

$$I_{ik} = I_k + \zeta_{Iik}, \tag{6.13}$$

$$S_{ik} = S_k + \gamma_{Zk}\,Z_i + \zeta_{Sik}, \tag{6.14}$$

where the assignment of an intervention has a differential effect (γ_{Zk}) on the growth rate of high compliers and low compliers. According to Equation 6.4, γ_{Zk} can be translated into the intervention effect at the last time point (i.e., $\gamma_{Zk} \times 3$).

Given that $K = 2$ and there are no covariates, the multinomial logit model can be simplified as

$$\begin{aligned} P(i \in h) &= \pi_{hi}, \\ P(i \in l) &= 1 - \pi_{hi} = \pi_{li}, \\ logit(\pi_{hi}) &= \beta_0, \end{aligned} \tag{6.15}$$

where π_{hi} denotes the probability of being a high complier, π_{li} denotes the probability of being a low complier, and β_0 represents a logit intercept that determines the ratio of high and low compliers. The true logit intercept is 0.0 (i.e., 50:50).

The model used for the CACE analysis based only on the observed variables (ANCOVA approach) can be expressed as

$$y_{i4} = \alpha_k + \lambda_x\, y_{i1} + \Gamma_{Zk}\, Z_i + \epsilon_{i4}, \tag{6.16}$$

where the differential effect of intervention (Γ_{Zk}) is defined based on the outcome measured at the last time point (y_{i4}) conditioning on the outcome measured at the first time point (y_{i1}). Here, the baseline outcome measure (y_{i1}) is considered as a covariate.

In the ANCOVA approach, the baseline outcome measure (y_{i1}) is also used as a predictor of compliance. Treating y_{i1} as a covariate, the logit model can be described as

$$\begin{aligned} P(i \in h \mid y_{i1}) &= \pi_{hi}, \\ P(i \in l \mid y_{i1}) &= 1 - \pi_{hi} = \pi_{li}, \\ logit(\pi_{hi}) &= \beta_0 + \beta_1\, y_{1i}, \end{aligned} \tag{6.17}$$

where the logit coefficient β_1 shows the level of association between the baseline outcome measure and compliance behavior.

The simulation results presented in Table 6.1 are based on 500 replications with a sample size of 300. Coverage is defined as the proportion of replications out of 500 replications where the true intervention effects for high and low compliers are covered by the 95% confidence intervals of intervention effect estimates. Power is defined as the proportion of replications out of 500 replications where the intervention effect estimate is significantly different from zero ($\alpha = .05$).

It is demonstrated in Table 6.1 that both the ANCOVA and growth model approaches provide average intervention effect estimates with reasonable quality, considering that compliance information is missing for 50% of individuals (i.e., control condition individuals) and the sample size is fairly small (i.e., $N = 300$). Simulation results show that the quality

TABLE 6.1

CACE Analyses Using the Observed and Latent Variable Approaches: The Quality of Average Intervention Effect Estimates at the Last Time Point

	Observed Variable (ANCOVA) Approach		
Intervention Effect	*True Value*	*Avg Estimate*	*Avg SE*
High Complier (Γ_{Zh})	0.60	0.604	0.292
Low Complier (Γ_{Zl})	-0.30	-0.280	0.291
Coverage (High Complier)		**0.930**	
Coverage (Low Complier)		**0.932**	
Power (High Complier)		**0.560**	
Power (Low Complier)		**0.226**	
	Latent Variable (Growth Model) Approach		
Intervention Effect	*True Value*	*Avg Estimate*	*Avg SE*
High Complier ($\gamma_{Zh} \times 3$)	0.60	0.606	0.259
Low Complier ($\gamma_{Zl} \times 3$)	-0.30	-0.286	0.260
Coverage (High Complier)		**0.940**	
Coverage (Low Complier)		**0.938**	
Power (High Complier)		**0.657**	
Power (Low Complier)		**0.248**	

of estimates is close between the two models in terms of point estimates and standard errors, implying comparability of the two models. The growth model approach, however, shows slightly better point estimates and standard errors than the ANCOVA approach. Although the gap between the two approaches in point estimates and standard errors is not dramatic, it still results in a noticeable difference between the two methods in terms of statistical power to detect intervention effects (e.g., 0.657 vs. 0.560 for high compliers). In the ANCOVA approach, only the outcome measures at the first (y_1) and the last time point (y_4) are considered, and the outcome measures in between (y_2, y_3) are ignored. The loss in information may lead to a lower precision in the ANCOVA approach.

It has been demonstrated in previous research that average causal effects of interventions can be identified for more than one subpopulation with a satisfactory level of accuracy based on auxiliary information from covariates (Jo, in press-a). Simulation results shown in Table 6.1 show that growth trajectories (in a latent variable form) can provide auxiliary information that can be used for the same purpose. It is also shown that the growth model approach may improve precision of average causal effect estimates by handling measurement errors and by utilizing trajectory information.

THE JOHNS HOPKINS PUBLIC SCHOOL PREVENTIVE INTERVENTION STUDY

The Johns Hopkins Public School Preventive Intervention Study was conducted by the Johns Hopkins University Preventive Intervention Research Center (JHU PIRC) in 1993 to 1994 (Ialongo et al., 1999). Based on the life course/social field framework as described by Kellam and Rebok (1992), the Johns Hopkins PIRC preventive trial focused on successful adaptation to first grade as a means of improving social adaptational status over the life course. The study was designed to improve academic achievement and to reduce early behavioral problems of school children. Teachers and first-grade children were randomly assigned to intervention conditions. The control condition and the family-school partnership intervention condition are compared in this example. In the intervention condition, parents were asked to implement 66 take-home activities related to literacy and mathematics over a 6 month period. The intervention was provided over the first-grade school year (1993–1994), following a pretest assessment in the early fall. The intervention impact was assessed in the spring of first (6 months from the pretest) and second (18 months from the pretest) grades. In the spring of first grade, 91.3% completed assessments, and in the spring of second grade, 88.5% completed assessments.

A total sample size of 333 was analyzed after listwise deletion of cases that had missingness in covariates and outcome variables. The two major outcome measures in the JHU PIRC preventive trial were academic achievement (CTBS mathematics and reading test scores) and the TOCA-R score (Teacher Observation of Classroom Adaptation-Revised; Werthamer-Larsson, Kellam, & Wheeler, 1991). The TOCA-R is designed to assess each child's adequacy of performance on the core tasks in the classroom as rated by the teacher. Among various outcome measures, readiness to learn (or work) assessed in the spring of the second grade (18 months from the pretest) is used as the outcome in this example. In the JHU PIRC preventive trial, readiness to learn does not represent acquisition of prerequisite knowledge or skills, but rather, being ready to exert effort to reach academic excellence. The readiness to learn scale ranges from 1 to 6, and consists of TOCA-R items that measure whether a child completes assignments, puts forth effort, and works hard. Table 6.2 shows the sample statistics for the variables used in the analyses of this study.

Intent to treat analysis using the ANCOVA approach

Standard ITT analysis provides an overall average intervention effect estimate by comparing the outcome based on assignment of intervention, but ignoring the aspect of the receipt of the intervention. That is, it assumes that children of parents with a low compliance rate receive the same effects from the intervention as children of parents with a high compliance rate. Table 6.3 shows the results from the JHU PIRC preventive trial data analysis using the ITT analysis. In this analysis, the overall effect of the intervention is estimated based on the outcome measured at the last time point (Ready3). The outcome measured at the first time point (Ready1) is used as one of the covariates, and the outcome measured at the second time point (Ready2) is not considered in the analysis (ANCOVA approach).

There is a positive effect of the intervention on the level of children's readiness to learn (intervention effect = 0.316, effect size = 0.212). The effect size of the intervention was calculated by dividing the outcome difference in the intervention and the control condition means by the square root of the variance pooled across the control and intervention groups. In the ITT analysis, baseline readiness to learn (Ready1) and free lunch program were found to be significant predictors of the level of readiness to learn. Children had a higher level of readiness at the last time point if their baseline readiness level was higher, and a lower level of readiness if their SES background level was low.

TABLE 6.2

Johns Hopkins PIRC: Sample Statistics (N=333)

Variable	*Mean*	*SD*	*Description*
Z	0.52	0.50	Intervention assignment (0 = control, 1 = intervention)
Ready1	4.59	1.32	TOCA mean readiness at the pretest
Ready2	4.48	1.39	TOCA mean readiness 6 months from the pretest
Ready3	4.33	1.49	TOCA mean readiness 18 months from the pretest
Male	0.49	0.50	Student's gender (0 = female, 1 = male)
Lunch	0.60	0.49	Free lunch program (0 = no, 1 = yes)
Unemployed	0.14	0.34	Parent's employment status (0 = no, 1 = yes)
Married	0.47	0.50	Parent's marital status (0 = no, 1 = yes)
Limited Health	0.10	0.30	Parent limited by health problem (0 = no, 1 = yes)
Health	3.83	1.03	Parent's overall health (1 = poor, 5 = excellent)
Age	2.97	1.42	Parent's age in 5-year brackets

TABLE 6.3

Johns Hopkins PIRC: Intent to Treat Analysis

Parameter	*Estimate*	*SE*
Intervention effect	0.316	0.147
Ready3 Regressed on Covariates		
Ready1	0.505	0.060
Male	-0.243	0.143
Free Lunch	-0.410	0.156
Unemployed	-0.348	0.236
Married	-0.288	0.147
Limited Health	-0.074	0.236
Health	-0.062	0.077
Age	-0.037	0.050
Intercept	2.742	0.503
$\sigma_3{}^2$	1.636	0.112

CACE analysis using the ANCOVA approach

In the ITT analysis, intervention effect may be underestimated for high compliers due to the inclusion of low compliers who might not have been exposed enough to benefit from the intervention. In this situation, the possible bias can be avoided by taking into account the difference between the two subpopulations in the analysis. Table 6.4 shows the results from the CACE analysis, where the differential effect of intervention is estimated for high compliers and low compliers. As in the ITT analysis, intervention effect is estimated based only on observed variables (ANCOVA approach). The same set of covariates used in the ITT analysis are used as predictors of the outcome, and also as predictors of compliance. The model used for this CACE analysis is the same as the model described in Equations 6.16 and 6.17, with the exception that more covariates are included in this example in addition to the baseline outcome measure.

Table 6.4 shows that the intervention had a positive impact on the

TABLE 6.4

Johns Hopkins PIRC: CACE Analysis Using the Observed Variable (ANCOVA) Approach

Parameter	*Estimate*	*SE*
Intervention Effect		
High Complier (Γ_{Zh})	0.477	0.221
Low Complier (Γ_{Zl})	-0.103	0.530
Ready3 Regressed on Covariates		
Ready0	0.524	0.060
Male	-0.246	0.145
Free Lunch	-0.419	0.157
Unemployed	-0.355	0.238
Married	-0.276	0.147
Limited Health	-0.078	0.243
Health	-0.059	0.079
Age	-0.048	0.051
High Comp Intercept (α_h)	2.523	0.523
Low Comp Intercept (α_l)	3.093	0.723
${\sigma_3}^2$	1.606	0.118
c Regressed on Covariates (High vs. Low Compliers)		
Ready0	0.376	0.130
Male	-0.021	0.350
Free Lunch	-0.235	0.385
Unemployed	-0.162	0.492
Married	0.095	0.347
Limited Health	-0.266	0.720
Health	0.009	0.191
Age	-0.221	0.109
Logit Intercept (β_0)	0.059	1.116

level of readiness for children with parents with a high compliance rate (intervention effect = 0.477, effect size = 0.320), and the magnitude of the effect was larger than that of the overall effect in the ITT method. It also shows a slightly negative effect of the intervention for children of parents with a low compliance rate, but the magnitude of the effect is very small and insignificant (intervention effect = -0.103, effect size = -0.069). Effect size was calculated based on a pooled standard deviation as in the ITT analysis. This approach was chosen for easier comparison across different estimation methods. In this analysis, baseline readiness to learn (Ready1) and free lunch program were found to be significant predictors of the level of readiness to learn. Children had a higher level of readiness at the last time point if their baseline readiness level was higher, and a lower level of readiness if their SES background level was low. Initial level of child's readiness and parent's age were found to be significant predictors of parent's compliance behavior. Parents complied more if the child's baseline readiness level was higher. Younger parents also complied more.

For the CACE analysis shown in Table 6.4, individuals were dichotomized into either the low or the high complier category based on the level of completeness in home learning activities. For easier comparison, the same cutpoint is used as that in the CACE analysis using the latent variable approach that is shown in a later section. For illustration purposes, compliance was dichotomized in this example; but note that sensitivity of the CACE estimate to different thresholds needs to be carefully examined in practice (West & Sagarin, 2000). The following section shows how the cutpoint was decided in this study for CACE analyses.

Exploratory growth mixture analysis

This section examines the possibility of using exploratory growth mixture analysis as a data-mining tool that precedes CACE analysis using mixture and growth mixture models. In randomized intervention trials, the intervention condition often includes many sessions, or doses. One way to model compliance behavior in this situation is to treat compliance as a continuous variable. Holland (1988) proposed ALICE (additive linearly constant effects) model, where the effect of intervention is estimated based on continuous compliance. The ALICE model requires several strong assumptions, which often limits the applicability of the model in practice. For example, it is assumed in the ALICE model that the effect of intervention linearly increases as the level of compliance increases. In the JHU PIRC preventive trial, there is a large variation in completed number of intervention activities (range 0 to 66), and children may not get any benefit from the intervention unless parents complete a sufficient number of activities. Over reporting of compliance level is also expected, because parents self-report their level of completion in intervention activities. In this situation, the intervention may not show any desirable effects unless parents report a quite high level of compliance. Another way to model compliance behavior in this situation is to use the dose-response curve approach (Efron & Feldman, 1991), where the effect of intervention can be estimated without assuming a linear relationship between intervention effects and compliance. However, this approach requires successful double-blind experiments, which are not often applicable especially in psychosocial intervention trials. The third way to model compliance behavior in this situation is to treat compliance as a categorical variable without assuming linearity. The difficulty of this approach is in deciding the appropriate number of categories and thresholds that separate individuals into different compliance categories.

The current study takes the third approach in analyzing the Johns Hopkins PIRC preventive trial data, and shows that exploratory growth mixture analysis can be useful in determining cutpoints at the planning stage of CACE analysis. To estimate the differential effect of the intervention for those who completed enough activities and for those who did not, the compliance measure is dichotomized in this example. Exploratory growth mixture analysis is conducted for control group individuals, which provides the information about the trajectory shape and the proportion of subgroups in the absence of intervention (Muthén et al., in press). The confirmatory mixture analysis (i.e., CACE analysis) following the exploratory analysis is based on the idea that subpopulations that are already different in the absence of intervention will be more likely to differ in terms of compliance behavior. Consequently, it is also expected that the effect of intervention will differ for these heterogeneous subpopulations.

The model used for exploratory growth mixture analysis is the same as the model described in Equations 6.1, 6.2, 6.3, and 6.5 except that $\gamma_{Zk} Z_i$ is removed from Equation 6.2 in the exploratory analysis. The same covariates used in CACE analyses in the previous section and in the following section are used in the exploratory growth mixture analysis. However, note that the model used for exploratory analysis is significantly different from the model used for mixture and growth mixture CACE analyses because the intervention group is not included in the model and potential level of compliance is not considered in estimating class membership of individuals. Figure 6.1 shows estimated trajectories suggested by two-class exploratory growth mixture analysis for the control group.

Figure 6.1 shows that the level of readiness decreases over time for the majority of children (69.2%), whose baseline readiness level is high. It also shows that the level of readiness increases over time for the other class

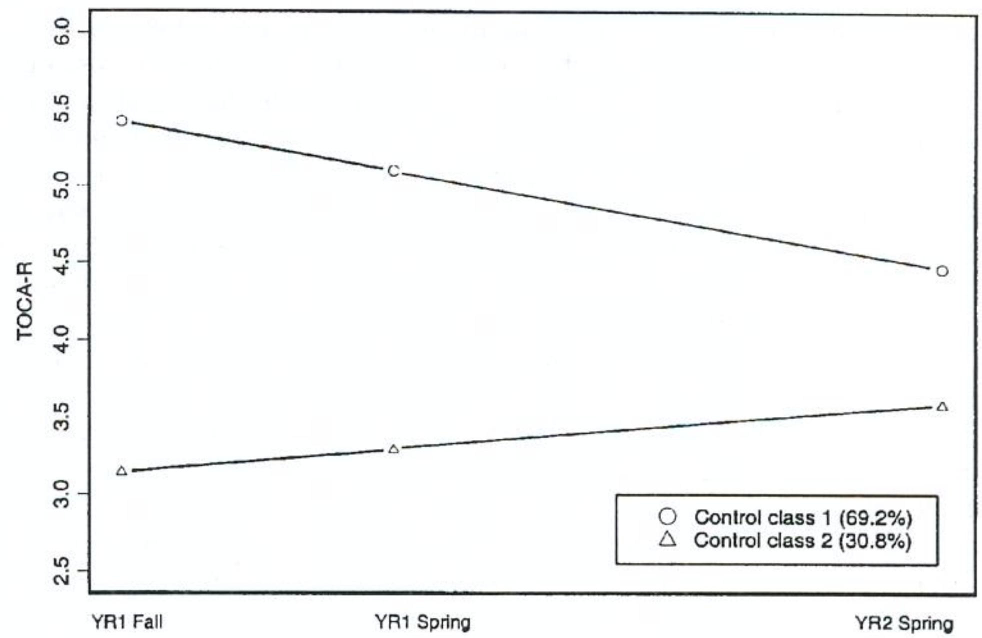

FIG. 6.1. Estimated mean curves of readiness to learn in the control group using exploratory growth mixture analysis.

of children (30.8%), whose baseline readiness level is low. Based on the proportion of subpopulations from the exploratory analysis, parents who completed 35 or more of the take-home learning activities were categorized as high compliers (71% of parents), and parents who completed fewer than 35 take-home learning activities were categorized as low compliers (29% of parents). Four parents did not comply at all and were included in the low complier category in this example. Parents in the control condition could not be dichotomized because their compliance information was missing. Figure 6.2 shows observed mean curves of readiness to learn based on this dichotomization.

CACE analysis using the growth model approach

This section demonstrates the estimation of intervention effects using the growth mixture modeling approach, where the effect of intervention is defined based on a trend or a growth trajectory of individuals. The growth mixture model used for CACE analysis is the same as the model described in Equations 6.1, 6.2, 6.3, and 6.5. The same covariates used in the CACE analysis using the ANCOVA approach and the exploratory growth mixture analysis are used. Based on exploratory growth mixture analysis of the control group and observed mean curves shown in Figure 6.2, linear

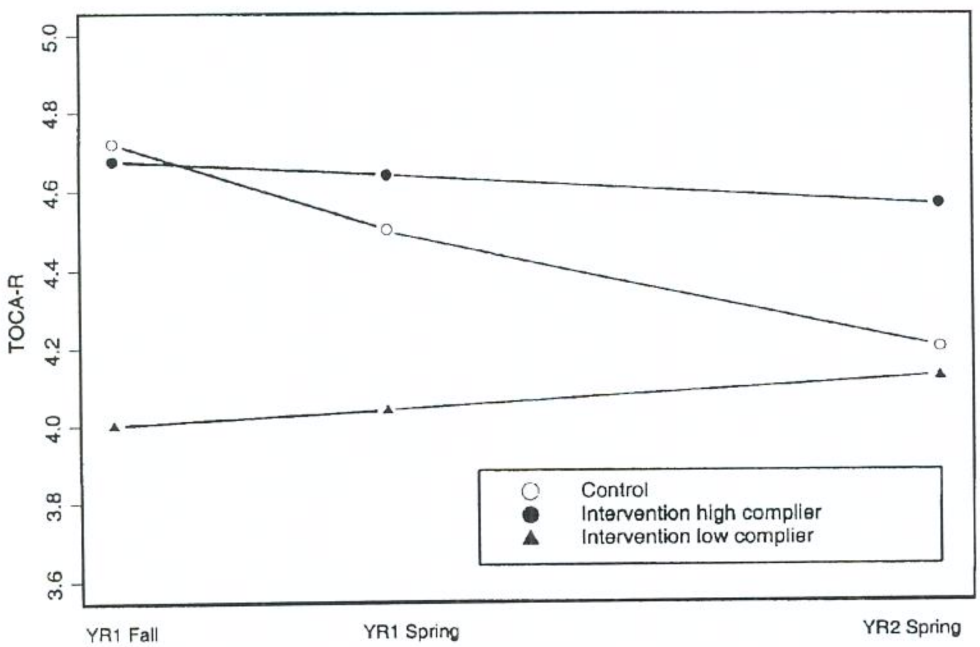

FIG. 6.2. Observed mean curves of readiness to learn.

trajectory was found to be appropriate for the CACE analysis using growth model approach. In the JHU PIRC preventive trial example, the outcome was measured in fall of the first grade, spring of the first grade, and spring of the second grade (Ready1, Ready2, Ready3). Given the distances between time points, the time scores used to capture a linear trend over time are 0, 1, and 3. Therefore, the average causal effect of intervention assignment at the last time point is defined as $\gamma_{Zh} \times 3$ for high compliers and $\gamma_{Zl} \times 3$ for low compliers (see Equation 6.4). Table 6.5 shows the results from the CACE analysis using the growth model approach.

Table 6.5 shows that the intervention had a positive impact on readiness of children if parents showed a high compliance rate (intervention effect = 0.477, effect size = 0.320). For children with highly complying parents, the level of readiness to learn decreases significantly less compared to that of control condition children with parents who could have been high compliers if they have had been assigned to the intervention condition. The intervention effect for high compliers in the CACE analysis using the growth model approach has the same magnitude as in the CACE analysis using the ANCOVA approach (see Table 6.4), but the confidence interval is slightly tighter than in the CACE analysis using the ANCOVA approach. Table 6.5 also shows a slightly negative but insignificant effect of the intervention for children of parents with a low compliance rate (intervention effect =

TABLE 6.5
Johns Hopkins PIRC: CACE Analysis Using the Latent Variable (Growth Model) Approach

Parameter	*Estimate*	*SE*
Intervention Effect		
High Complier ($\gamma_{Zh} \times 3$)	0.477	0.186
Low Complier ($\gamma_{Zl} \times 3$)	-0.150	0.393
Initial Status Regressed on Covariates		
Male	-0.259	0.134
Free Lunch	-0.244	0.146
Unemployed	-0.068	0.200
Married	0.076	0.138
Limited Health	0.368	0.254
Health	0.041	0.078
Age	0.058	0.047
High Comp Intercept (I_h)	4.682	0.420
Low Comp Intercept (I_l)	3.908	0.427
ψ_{Ih}	0.936	0.139
ψ_{Il}	1.608	0.253
Growth Rate Regressed on Covariates		
Male	-0.033	0.054
Free Lunch	-0.092	0.057
Unemployed	-0.108	0.091
Married	-0.105	0.055
Limited Health	-0.093	0.101
Health	-0.025	0.031
Age	-0.026	0.019
High Comp Intercept (S_h)	0.118	0.158
Low Comp Intercept (S_l)	0.424	0.208
ψ_{Sh}	0.064	0.037
ψ_{Sl}	0.151	0.055
ψ_{ISh}	-0.048	0.051
ψ_{ISl}	-0.227	0.106
${\sigma_1}^2$	0.444	0.118
${\sigma_2}^2$	0.708	0.080
${\sigma_3}^2$	0.717	0.216
c Regressed on Covariates (High vs. Low Compliers)		
Male	-0.175	0.324
Free Lunch	-0.388	0.352
Unemployed	-0.217	0.457
Married	0.102	0.331
Limited Health	0.042	0.621
Health	0.029	0.171
Age	-0.205	0.106
Logit Intercept (β_0)	1.774	0.874

-0.150, effect size = -0.101).

The parameterization used in the growth model approach adds more flexibility in the interpretation of the CACE analysis than that used in the ANCOVA approach because initial status (I) and growth rate (S) are separated. For example, the influence of background variables can be estimated separately for initial level of readiness and change of readiness. In the CACE analysis using the ANCOVA approach, child's gender and free lunch program were found to be significant predictors of the outcome. However, these variables are not significant predictors when initial status (I) and growth rate (S) are separated as shown in Table 6.5. It also shows in CACE analysis using the growth model approach that low compliers have more variation in initial status and growth rate conditioning on covariates. In the high compliance category, initial status and growth rate show very low correlation conditioning on covariates ($\psi_{ISh} = -0.048$). However, in the low compliance category, initial status and growth rate are negatively correlated conditioning on covariates ($\psi_{ISl} = -0.227$). In addition to flexibility in modeling, another advantage of the growth model approach is that it utilizes not only covariates but also trajectory information to identify class membership and to increase efficiency in estimating the differential effect of intervention.

Figure 6.3 shows estimated mean readiness curves over time based on results in Table 6.5. Estimated mean outcomes can be calculated using Equations 6.1, 6.2 and 6.3 and weighted covariate means based on posterior class probability of each individual. This figure shows how readiness to learn changed over time depending on parents' compliance level and intervention assignment. It shows that highly complying parents' children had a higher level of readiness at the first grade, but the level could decrease to a point even lower than that of less involved parents' children by the second grade unless the intervention was given.

By comparing the mean trajectories of the control group in Fig. 6.3 to those in Fig. 6.1, it can be learned how closely subpopulations derived by exploratory and confirmatory growth mixture analyses are related. Mean trajectories in Figs. 6.1 and 6.3 show similarity in the sense that the level of readiness decreases over time for the majority of children (those with high baseline readiness), and the level of readiness increases over time for the other class of children (those with low baseline readiness). Mean trajectories in Figs. 6.1 and 6.3 also show discrepancy in the sense that trajectories in Fig. 6.1 have a larger difference at the initial point and a smaller difference at the last time point than those in Fig. 6.3. The disagreement is not surprising because the model used for exploratory analysis does not consider potential level of compliance, whereas CACE (confirmatory) analysis does. However, information from exploratory analysis is still valuable in deciding

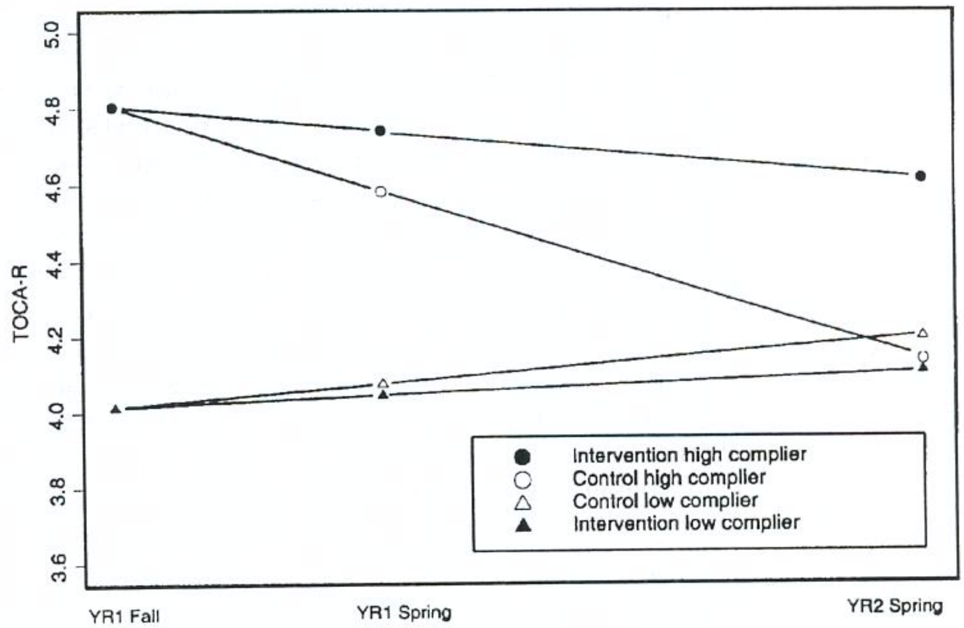

FIG. 6.3. Estimated mean curves of readiness to learn: CACE analysis using the growth model approach.

the cutpoint that is essential for CACE analysis, given that there are no established methods that can determine an optimal cutpoint. Without information from exploratory growth mixture analysis, one may simply choose to categorize 50% of individuals into the high complier category and 50% of individuals into the low complier category. In the JHU PIRC preventive trial example, the CACE analysis based on the cutpoint from exploratory growth mixture analysis was found to be substantially better than the CACE analysis based on the simple categorization (i.e., 50:50) in terms of model fit, precision of intervention effect estimates, and precision in classification of individuals into different compliance categories.

CONCLUSION

This study demonstrated the estimation of differential average intervention effects at varying levels of compliance in a growth mixture modeling framework, where the effect of the intervention is defined based on a trend or a growth trajectory of individuals. It was demonstrated in simulation studies that the quality of intervention effect estimates in ANCOVA and growth model approaches is very close, implying comparability of the two approaches. The growth model approach, however, showed slightly better point estimates and standard errors than did the ANCOVA approach.

Although the gap between the two approaches in point estimates and standard errors was not dramatic, it still resulted in a noticeable difference between the two methods in terms of statistical power to detect intervention effects. In the growth model approach, precision of average causal effect estimates can be improved by handling measurement errors and by utilizing trajectory information. In contrast, the ANCOVA approach only considers the outcome measured at the first and the last time point, and ignores the outcome measured in between. The loss in information may lead to lower precision in the ANCOVA approach.

In the JHU PIRC preventive trial example shown in this study, the differential effect of intervention was estimated through CACE analysis using ANCOVA and growth model approaches. CACE analyses using ANCOVA and growth mixture approaches showed a larger effect of intervention for high compliers compared to the overall effect in the ITT method. The results were also compared between CACE analyses using the ANCOVA and growth model approaches. In line with simulation study results, it was shown in this example that ANCOVA and growth model approaches have close intervention effect estimates, but CACE analysis using the growth model approach showed a slightly tighter confidence interval than CACE analysis using the ANCOVA approach. It was also demonstrated that the parameterization of the growth model approach adds more flexibility in modeling and provides richer information than that of the ANCOVA approach.

In the JHU PIRC preventive trial example shown in this study, individuals were classified into two groups, and CACE models were identified based on various covariates and growth trajectories. The exclusion restriction could not be assumed in this example, because low compliers were also exposed to the intervention. The intervention might have had a weaker impact on low compliers, but it cannot be guaranteed that the intervention had no effect at all. Without assuming the exclusion restriction, the identifiability and the quality of CACE estimation relies on auxiliary information (Hirano et al., 2000; Jo, in press-a). Given that, it is desirable to use multiple sources of information to improve accuracy and efficiency in the estimation. In the JHU PIRC preventive trial example shown in this study, not only covariate information but also trajectory information was used to identify class membership and to increase efficiency in the estimation of differential intervention effects. Although previous research showed that it is possible to identify CACE models without assuming the exclusion restriction based on auxiliary information, very little is known about how this method should be applied in practice. More research is needed in this area to explore what kind of information and modeling approaches are more efficient and how stability of models should be checked in extended versions of CACE models.

This study also examined the possibility of using exploratory growth mixture analysis as a data-mining tool that precedes CACE analyses. The confirmatory mixture analysis (i.e., CACE analysis) following the exploratory mixture analysis is based on the idea that subpopulations that are already different in the absence of intervention will be more likely to differ in compliance behavior. How closely subpopulations derived by exploratory growth mixture analyses are related to subpopulations derived by CACE analysis varies in different situations. When the intervention condition includes many sessions, as in the JHU PIRC preventive trial, how individuals are categorized into different compliance classes is critical for CACE analysis. However, little is known about how to determine optimal cutpoints and number of cutpoints. Given that, exploratory growth mixture analysis can be useful at the planning stage of CACE analysis in the sense that it provides information about subpopulations that are heterogeneous in the absence of the intervention. Further research is needed in this area to establish a systematic way of connecting exploratory and confirmatory mixture analyses.

ACKNOWLEDGMENTS

This study was supported by National Institute on Alcohol Abuse and Alcoholism Grant K02 AA 00230 (Bengt Muthén, Principal Investigator) and National Institute of Mental Health Grant P50 MH38725 (Philip Leaf, Principal Investigator). We would like to thank Nicholas Ialongo and Sheppard Kellam for their data and insightful advice.

REFERENCES

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*, 444–455.

Bloom, H. S. (1984). Accounting for no-shows in experimental evaluation designs. *Evaluation Review, 8*, 225–246.

Dempster, A., Laird, N., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, 39,* (Series B), 1–38.

Efron, B., & Feldman, D. (1991). Compliance as an explanatory variable in clinical trials. *Journal of the American Statistical Association, 86*, 9–17.

Hirano, K., Imbens, G. W., Rubin, D. B., & Zhou, X. H. (2000). Assessing

the effect of an influenza vaccine in an encouragement design. *Biostatistics, 1*, 69–88.
Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*, 945–970.
Holland, P. W. (1988). Causal inference, path analysis, and recursive structural equation models. In C. Clogg (Ed), *Sociological methodology* (pp. 449–484). Washington, DC: American Sociological Association.
Ialongo, N. S., Werthamer, L., Kellam, S. G., Brown, C. H., Wang, S., & Lin, Y. (1999). Proximal impact of two first-grade preventive interventions on the early risk behaviors for later substance abuse, depression and antisocial behavior. *American Journal of Community Psychology, 27*, 599–642.
Imbens, G. W., & Angrist. J. (1994). Identification and estimation of local average treatment effects. *Econometrica, 62*, 467–476.
Imbens, G. W., & Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with non-compliance. *The Annals of Statistics, 25*, 305–327.
Jo, B. (in press-a). Estimating intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics.*
Jo, B. (in press-b). Model misspecification sensitivity analysis in estimating causal effects of interventions with noncompliance. *Statistics in Medicine.*
Jo, B. (2002). Statistical power in estimating causal effects of interventions with noncompliance. *Psychological Methods, 7*, 178–193.
Jo, B., & Muthén, B. O. (2001). Modeling of intervention effects with noncompliance: A latent variable modeling approach for randomized trials. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 57–87), Lawrence Erlbaum Associates.
Kellam, S. G., & Rebok, G. W. (1992). Building developmental and etiological theory through epidemiologically based preventive intervention trials. In J. McCord & R. E. Tremblay (Eds.), *Preventing anti-social behavior: Interventions from birth through adolescence* (pp. 162–195). New York: Guilford.
Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data.* New York: Wiley.
Little, R. J. A., & Yau, L. (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., & Krishnan, T. (1997). *The EM algorithm and extensions.* New York: Wiley Series.

McLachlan, G. J., & Peel, D. (2000). *Finite mixture models.* New York: Wiley Serires.

Muthén, B. O. (2001a). Latent variable mixture modeling. In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling* (pp. 1–33), Lawrence Erlbaum Associates.

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

Muthén, B. O., Brown, C. H., Khoo, S. T., Yang, C. C., & Jo, B. (1997). *General growth mixture modeling of latent trajectory classes: Perspectives and prospects.* Manuscript in preparation.

Muthén, B. O., Brown, C. H., Masyn, K., Jo, B., Khoo, S. T., Yang, C. C., Wang, C. P., Kellam, S. G., Carlin, J. B., & Liao, J. (in press). General growth mixture modeling for randomized preventive interventions. *Biostatistics.*

Muthén, L. K. & Muthén, B. O. (1998-2001). M*plus user's guide.* Los Angeles: Muthén & Muthén.

Muthén, B. O., & 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.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology, 66,* 688–701.

Rubin, D. B. (1978). Bayesian inference for causal effects: The role of randomization. *Annals of Statistics, 6,* 34–58.

Rubin, D. B. (1980). Discussion of "Randomization analysis of experimental data in the Fisher randomization test" by D. Basu. *Journal of the American Statistical Association, 75,* 591–593.

Rubin, D. B. (1990). Comment on "Neyman (1923) and causal inference in experiments and observational studies." *Statistical Science, 5,* 472–480.

Tanner, M. (1996). *Tools for statistical inference: Methods for the exploration of posterior distributions and likelihood functions.* New York: Springer.

Titterington, D. M., Smith, A. F. M., & Makov, U. E. (1985). *Statistical analysis of finite mixture distributions.* Chichester, England: Wiley.

Werthamer-Larsson, L., Kellam, S. G., & Wheeler, L. (1991). Effect of first-grade classroom environment on child shy behavior, aggressive behavior, and concentration problems. *American Journal of Community Psychology, 19*, 585–602.

West, S. G., & Sagarin, B. J. (2000). Participant selection and loss in randomized experiments. In L. Bickman (Ed.), *Research design* (pp. 117–154). Thousand Oaks, CA: Sage.