Satorra, A., & Bentler, P. M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *American Statistical Association 1988 proceedings of Business and Economics Sections* (pp. 308–313). Alexandria, VA: American Statistical Association.

Satorra, A., & Bentler, P. M. (1999). A scaled difference chi-square test statistic for moment structure analysis. University of California, Los Angeles: UCLA Statistics Series, No. 260 (www.stat.ucla.edu).

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sörbom, D. (1974). A general method for studying differences in factor means and factor structures between groups. *British Journal of Mathematical and Statistical Psychology, 27*, 229–239.

Tanaka, Y., Watadani, S., & Moon, S. H. (1991). Influence in covariance structure analysis: With an application to confirmatory factor analysis. *Communication in Statistics-Theory and Method, 20*, 3805–3821.

Tremblay, P. F., & Gardner, R. C. (1996). On the growth of structural equation modeling in psychological journals. *Structural Equation Modeling, 3*, 93–104.

Tyler, D. E. (1983). Robustness and efficiency properties of scatter matrices. *Biometrika, 70*, 411–420.

Yuan, K.-H., & Bentler, P. M. (1998a). Robust mean and covariance structure analysis. *British Journal of Mathematical and Statistical Psychology, 51*, 63–88.

Yuan, K.-H., & Bentler, P. M. (1998b). Structural equation modeling with robust covariances. *Sociological methodology, 28*, 363–396.

Yuan, K.-H., & Bentler, P. M. (2000a). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological methodology, 30*, 165–200.

Yuan, K.-H., & Bentler, P. M. (2000b). Robust mean and covariance structure analysis through iteratively reweighted least squares. *Psychometrika, 65*, 43–58.

Yuan, K.-H., Chan, W., & Bentler, P. M. (2000). Robust transformation with applications to structural equation modeling. *British Journal of Mathematical and Statistical Psychology, 53*, 31–50.

Yuan, K.-H., & Jennrich, R. I. (1998). Asymptotics of estimating equations under natural conditions. *Journal of Multivariate Analysis, 65*, 245–260.

Yuan, K.-H., Marshall, L., & Bentler, P. M. (1999). A unified approach to exploratory factor analysis with missing data, nonnormal data, and in the presence of outliers. Under review.

Yuan, K.-H., Marshall, L., & Weston, R. (1999). Cross-validation through downweighting influential cases in structural equation modeling. Under review.

#90

3

Modeling of Intervention Effects With Noncompliance: A Latent Variable Approach for Randomized Trials

Booil Jo
Bengt O. Muthén
University of California, Los Angeles

It is well known that experimental designs based on randomization are powerful in terms of statistical analysis and inference. However, the estimation of treatment effects can be biased even with successful randomization unless everyone complies with the given treatment. Noncompliance is not only an obstacle to fair statistical comparison between the treatment group and the control group, but also a major threat to obtaining power to detect intervention effects (Jo, 2000c). Depending on how noncompliance is dealt with in the estimation of treatment effects, different conclusions may be reached about the effect of the same intervention trial.

Figure 3.1 illustrates subgroups in the intervention trial based on treatment assignment and compliance. It is shown that belonging to the complier or noncomplier category is not randomized but chosen by individuals, whereas the assignment to treatment or control condition is randomized. In the treatment condition, compliance behavior is actually observed and individuals can be categorized into either the *complier* or *noncomplier* category. In the control condition, compliance behavior cannot be observed because treatment is never offered. Therefore, individuals in the control condition are potentially either complier or noncomplier, but cannot be categorized based on observed compliance behavior. Potential compliers are individuals in the control condition who would comply

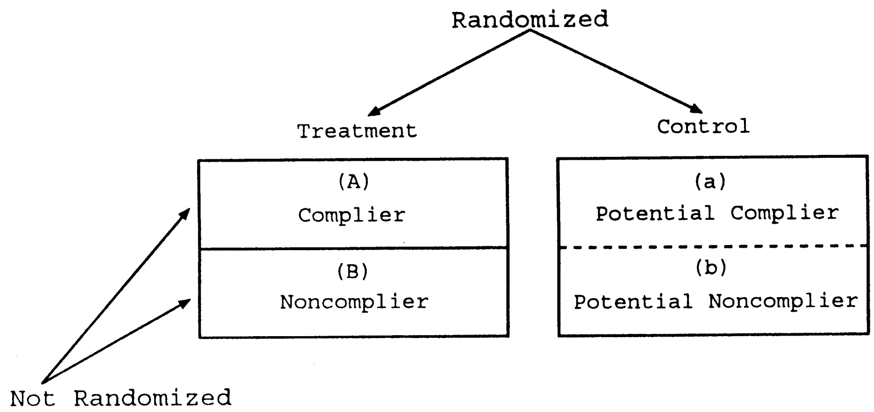

FIG. 3.1. Randomization and compliance.

with the treatment if it had been offered. Potential noncompliers are individuals in the control condition who would not comply with the treatment even if it had been offered.

Intent to Treat (ITT) analysis is a standard way to estimate treatment effects in randomized experimental designs. In this method, average outcomes are compared by randomized groups ignoring compliance status information. In other words, the treatment effect is estimated assuming that every subject in the treatment condition actually received the treatment. It is shown in Fig. 3.1 that the treatment $(A+B)$ and the control $(a+b)$ groups are statistically comparable in this method because both groups consist of both compliers and noncompliers. However, if only compliers are the targeted subpopulation of interest, there is a possible bias in the estimation of treatment effects by including noncompliers in the analysis.

As-treated analysis is another commonly used method to estimate intervention effects in the presence of noncompliance. This method focuses on the receipt of the treatment, but ignores the fact that compliance behavior is not randomized but chosen by individuals, and the characteristics of compliers are often different from those of the rest. For example, people with higher motivation or a special interest in the treatment are more likely to participate in that treatment. This method presents an unfair statistical comparison between groups, by comparing recipients (A) in the treatment group with nonrecipients $(B+a+b)$ both in the treatment and the control group.

When compliers are the targeted subpopulation of interest, there is a possible bias in the estimation of treatment effects in the presence of noncompliance both

in ITT and as-treated analysis. To counter this unfair comparison, the possibility of estimating causal effects of the treatment only for the individuals who actually received the treatment has been explored under the label Complier Average Causal Effect (CACE; Angrist, Imbens, & Rubin, 1996; Bloom, 1984; Imbens & Rubin, 1997; Little & Yau, 1998) estimation. This method not only provides the estimation of treatment effects only for compliers, but also presents a fair statistical comparison by comparing the compliers (A) in the treatment group to the potential compliers (a) in the control group (see Fig. 3.1).

The CACE estimation method has been applied using several approaches. The major technical difficulty involved in CACE approaches is that the compliance status of the individuals in the control condition is unknown. The unknown compliance status in the control group makes it difficult to differentiate effects of the treatment based on compliance status. One way to solve this problem is to use the instrumental variable (IV) approach, where treatment effect estimates are adjusted by considering the proportion of noncompliers (Bloom, 1984). More recently, a refined form of the IV approach with clear underlying assumptions has been proposed (Angrist, Imbens, & Rubin, 1996). A more efficient way to solve this problem is to identify potential compliance status of the control group individuals so that average outcomes can be directly compared based on randomization. This method has been demonstrated through a Bayesian approach that combines the use of EM and data-augmentation algorithms (Imbens & Rubin, 1997) and the maximum-likelihood estimation method using the EM algorithm (Little & Yau, 1998). The idea of CACE made dramatic progress in the estimation of treatment effects in the presence of noncompliance. By introducing Bayesian inferential methods and missing data techniques, this approach opened the possibility for more flexible model-based estimation of treatment effects.

Structural equation modeling has potential for flexible CACE modeling. However, the exploration possibility of CACE modeling in this area is limited within the conventional framework. Although the unknown compliance status in the control group can be naturally seen as a missing data problem in general, subgroups of individuals based on compliance status can be better understood as a latent variable in the structural equation modeling framework. The systematic role of compliance categories distiguishes latent membership from missing data in outcome measures. That is, individuals in different compliance categories can be seen as finite mixtures (Titterington, Smith, & Makov, 1985) of subpopulations that might have separate distributions and different model parameters. If the compliance status is known for everybody, this problem can be solved using the multiple-group approach in conventional structural equation modeling. Because the group membership is unknown for individuals in the control group, this problem cannot be solved unless discrete latent variables can be included in the model.

The current study demonstrates that the problem of noncompliance can be dealt with in a broader framework of structural equation modeling by looking at compliance status as a categorical latent variable, and also demonstartes the flexibility of CACE modeling in this framework. To demonstrate how the latent variable approach works in dealing with compliance information in various situations, the Job Search Intervention Study for unemployed workers (Vinokur, Price, & Schul, 1995; Vinokur & Schul, 1997), the Study of Vitamin Supplement Effect on Survival Rates in young children (Imbens & Rubin, 1997; Sommer et al., 1986; Sommer & Zeger, 1991), and the Johns Hopkins Public School Preventive Intervention Study (Ialongo et al., 1999) are used as examples.

This chapter is organized as follows. First, it defines model assumptions and the estimation method using the ML-EM algorithm. Second, it demonstrates CACE estimation with a single continuous outcome, with results compared to those from the ITT approach. Third, it demonstrates CACE estimation with a single categorical outcome, with results compared to those from the ITT approach. Fourth, it demonstrates CACE estimation with multiple outcomes, with results compared to those from CACE estimation using a single outcome measure. Fifth, it demonstrates growth mixture CACE estimation using repeated outcome measures with a trend. Results are compared to those from CACE estimation using a single outcome measure. The chapter concludes with discussion.

CACE ESTIMATION IN THE LATENT VARIABLE MODELING FRAMEWORK

Model Assumptions

The common purpose of the models used in this study is to estimate the treatment effect for the compliers (CACE) and to draw causal inference about this treatment effect through experimental designs based on randomization. In line with Rubin's causal model, there are some general assumptions required to be able to make causal inference. 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). The assumption of potential exposability (Holland, 1988) implies that the nature of the treatment should be alterable so that individuals have the possibility of exposure to either condition, although they cannot be exposed to the treatment and the control condition at the same time. When this basic assumption is satisfied, 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. The models used to analyze compliance in this study assume randomization and SUTVA in line with Rubin's causal model.

Assume the simplest experimental setting where there is only one outcome measure (y), the treatment assignment (T) is binary (1 = treatment, 0 = control), and the treatment received (D) has only two levels (1 = received, 0 = not received). By classifying the behavior types of the subjects based on combinations of T and D, four types of subpopulations can be defined. These definitions are based on the individual level, which is possible because of the assumption of potential exposability. An individual i cannot be exposed to the treatment ($T_i = 1$) and the control condition ($T_i = 0$) at the same time, but has the possibility of exposure to either condition.

Angrist et al. (1996) labeled the four categories as *complier*, *never-taker*, *defier*, and *always-taker*. Compliers are subjects who do what they are assigned to do ($D_i = 1 | T_i = 1$, and $D_i = 0 | T_i = 0$). Never-takers are subjects who do not receive the treatment even if they are assigned to the treatment condition ($D_i = 0 | T_i = 1$, and $D_i = 0 | T_i = 0$). Defiers are the subjects who do the opposite of what they are assigned to ($D_i = 0 | T_i = 1$, and $D_i = 1 | T_i = 0$). Always-takers are the subjects who always receive the treatment no matter which condition they are assigned to do ($D_i = 1 | T_i = 1$, and $D_i = 1 | T_i = 0$).

Among these four kinds of possible compliance behaviors, the current study focuses on compliers and never-takers. That is, it is assumed that there are neither defiers nor always-takers. This is a stronger assumption than monotonicity (Imbens & Angrist, 1994) in the instrumental variable approach, where it is assumed that there are no defiers. Although defiers and always-takers are also possible compliance behaviors, the existence of never-takers is a more commonly seen problem. In examples shown in this study, subjects were not allowed to choose a different treatment condition than the one to which they were assigned. For never-takers, it is assumed that the outcome is independent of the treatment assignment (the exclusion restriction; Angrist et al., 1996), implying no assignment effects of the treatment. Based on these assumptions (randomization, SUTVA, monotonicity, no always-takers, and the exclusion restriction), two kinds of subpopulations can be defined: never-takers and compliers. For simplicity, never-takers are labeled as noncompliers in this chapter.

CACE Estimation Using ML-EM

The randomization in the assignment of treatment condition provides the basis for identification in CACE models. In addition to the equality in the parameter values based on random assignment assumption, the observed compliance status among treatment group individuals (training data) also plays a key role in the estimation of the treatment effect for the compliers (CACE).

Consider a single outcome variable y_{ik} for individual i within latent class k,

$$y_{ik} = \alpha_k + \Gamma_{Tk} T_i + \varepsilon_{ik}, \tag{1}$$

where latent categorical variable c has K levels of compliance status ($k = 1, 2, \ldots, K$). c represents observed compliance status in the treatment group and latent compliance status in the control group. $c_i = (c_{i1}, c_{i2}, \ldots, c_{ik})$ has a multinomial distribution, where $c_{ik} = 1$ 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 (Muthén et al., 1997; Titterington, 1985). In finite mixture modeling, the number of mixture components is assumed to be known and fixed. For example, $K = 2$ in examples shown in this study ($k = 1$ for compliers, $k = 2$ for noncompliers). ε_{ik} represents the normally distributed residual with zero mean independent of treatment assignment T (1 = treatment, 0 = control). Let $V(\varepsilon_{ik}) = \sigma_k^2$ be the residual variance within compliance class k. α_k is the mean for the control group within latent class k, and Γ_{Tk} is the intervention effect within latent class k. The parameters of interest in the CACE model are α_k, Γ_{Tk}, σ_k^2, and the proportion of the population from component k with $\Sigma_{k=1}^{K} \pi_k = 1$. The proportion of compliers is π_1, and the proportion of noncompliers is $1 - \pi_1 = \pi_2$.

The identifiability of the model can be shown by solving for these parameters in terms of the population quantities that have observable counterparts in the form of consistent estimates. As a first step, π_1 is directly identified as the observed proportion of compliers in the treatment condition $P(k = 1)$. The remaining parameters α_k and Γ_{Tk} are identified based on observed means and π_k.

Based on Eq. (1), the parameters that represent average treatment effects for compliers and noncompliers are defined as

$$\alpha_1 + \Gamma_{T1} - (\alpha_1 + 0) = \Gamma_{T1} = CACE \tag{2}$$

$$\alpha_2 + \Gamma_{T2} - (\alpha_2 + 0) = \Gamma_{T2}, \tag{3}$$

whereas the unknown control group means for compliers and noncompliers are

$$\mu_{C,k=1} = \alpha_1, \tag{4}$$

$$\mu_{C,k=2} = \alpha_2, \tag{5}$$

the treatment group means are

$$\mu_{T,k=1} = \alpha_1 + \Gamma_{T1}, \tag{6}$$

$$\mu_{T,k=2} = \alpha_2 + \Gamma_{T2}, \tag{7}$$

and the overall control group mean is

$$\mu_C = \pi_1 \mu_{C,k=1} + \pi_2 \mu_{C,k=2}. \tag{8}$$

Because $\Gamma_{T2} = 0$ under the exclusion restriction assumption, α_2 is directly identified from equation (7) as

$$\alpha_2 = \mu_{T,k=2}. \tag{9}$$

From Eqs. (4), (8), and (9), α_1 can then be expressed in terms of known quantities as

$$\alpha_1 = (\mu_C - \pi_2 \mu_{T,k=2})/\pi_1. \tag{10}$$

From Eqs. (6) and (10), the average treatment effect for compliers can be expressed in terms of known quantities:

$$CACE = \Gamma_{T1} = \mu_{T,k=1} - (\mu_C - \pi_2 \mu_{T,k=2})/\pi_1. \tag{11}$$

The parameters σ_1^2 and σ_2^2 can then be identified from the mixture distribution of y (Eq. [1]). Because variances are not involved in the identification of Γ_{T1} as shown earlier, CACE models can be identified in the same way (Eq. [2]–[11]) when the outcome measure is categorical.

A single binary outcome variable u_{ik} for individual i within latent class k can be defined in a logit form as

$$logit(\tau_{ik}) = \alpha_{uk} + \Gamma_{Tk} T_i, \tag{12}$$

where $\tau_{ik} = P(u_{ik} = 1 \mid c_{ik} = 1)$. α_{uk} represents the intercept in the logistic regression of u on T within compliance class k. Γ_{T1} can be defined as the treatment effect for compliers as in the CACE model with a continuous outcome measure.

This study also demonstrates CACE estimation in the random coefficient growth mixture modeling framework. The growth mixture CACE model can be expressed using a two-level formulation. Consider a single outcome variable y for individual i at time point h within compliance class k,

$$y_{ihk} = I_{ik} + S_{ik} h + \varepsilon_{ihk}, \tag{13}$$

where ε_{ihk} represents a vector of normally distributed residuals with zero mean independent of other variables in the model. Let $V(\varepsilon_{ihk}) = \sigma_{hk}^2$. I_{ik} and S_{ik} are individually varying continuous latent variables representing initial level of outcome and growth rate (slope), respectively. The time scores h are $0, 1, 2, \ldots, H$, representing linear growth over time, which may be fixed at different values depending on the distance between the measuring points. Individual variation in parameters I_{ik} and S_{ik} within compliance class k is specified in the second level as

$$I_{ik} = I_k + \zeta_{Iik}, \quad (14)$$

$$S_{ik} = S_k + \Gamma_{Tk} T_i + \zeta_{Sik}. \quad (15)$$

In Eqs. (14) and (15), I_k and S_k represent intercept parameters of initial status and slope for each compliance class k. ζ_{Iik} and ζ_{Sik} can differ at different levels of compliance status, but the common residual variances $V(\zeta_{Iik}) = \psi_I$ and $V(\zeta_{Sik}) = \psi_S$ are used across different compliance classes for simplicity of illustration in the examples shown for this study. Based on randomization, initial status is not regressed on T_i but growth rate (slope) is regressed on T_i. Γ_{Tk} represents a mean shift in the slope when subject i belongs to the treatment condition and is allowed to vary across different compliance status. Γ_{T1} can be identified in the same way as in the estimation of CACE using a single outcome measure. The difference is that the intervention effect is identified based on means of growth rate (latent variable) instead of observed outcome means. In a growth modeling framework, treatment effects can be defined either as the difference between treatment and control conditions in the growth rate or as the difference between treatment and control conditions in the oucome measure at the final time point (Muthén & Curran, 1997). The second definition is used in the study for easier comparison between an ANCOVA approach using univariate outcome and growth mixture CACE modeling. Based on Eqs. (13), (14), and (15), the average treatment effects for compliers (CACE) can be defined at the last time point as

$$CACE = \Gamma_{T1} \times H \quad (16)$$

When covariates are present, the information carried by the covariates influences the CACE model in two ways. First, the precision in the regression of y (or η) on T is affected by inclusion of covariates (e.g., ANCOVA). Second, the class probability π_i is allowed to vary as a function of covariates. The logistic regression model of c on a vector of covariates x is decribed in a logit form as

$$logit(\pi_{1i}) = \alpha_c + \beta_c x_i, \quad (17)$$

where π_{1i} denotes the probability of being a complier. Because it is assumed that the treatment assignment is random, π_1 is the same for the control and treatment groups. The logistic regression of compliance status also provides information about the characteristics of the compliers.

The maximum likelihood estimation method using the EM algorithm (Dempster, Laird, & Rubin, 1977; McLachlan & Krishnan, 1997; Tanner, 1996) is employed in the current study to estimate the unknown compliance status of each subject in the control condition and to estimate average treatment effects for compliers.

Consider the sampling distribution of y and x from the mixture of k components

$$g(y, x \mid \theta, \pi) = \sum_{k=1}^{K} \pi_k f(y, x \mid \theta_k), \quad (18)$$

where y and x represent observed data, θ 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 π is the parameter that determines the distribution of c. The observed data log likelihood is

$$Log\, L = \sum_{i=1}^{n} log[y_i \mid x_i]. \quad (19)$$

Given the proposed CACE model in the presence of both covariates (x) and continuous latent variables (η), the complete data log likelihood can be written as

$$Log\, L_c = \sum_{i=1}^{n} (log[c_i|x_i] + log[\eta_i|c_i, x_i] + log[y_i|c_i, \eta_i, x_i), \quad (20)$$

where

$$\sum_{i=1}^{n} log[c_i|x_i] = \sum_{i=1}^{n} \sum_{k=1}^{K} c_{ik}\, log\, \pi_{ik}. \quad (21)$$

In Eqs. (20) and (21), c represents categorical latent compliance class, and η represents continuous latent growth factors (e.g., I and S).

Maximum likelihood estimation using the EM algorithm maximizes the expected complete data log likelihood shown in Eq. (20). In maximizing the expected complete data log likelihood in Eq. (20), the E step computes the expected values of the complete data sufficient statistics given data and current parameter estimates. c is considered as missing data in this step. The conditional distribution of c given the observed data and the current value of model parameter estimates θ'' is given by

$$f(c \mid y, x, \theta) = \prod_{i=1}^{n} f(c_i \mid y_i, x_i, \theta). \quad (22)$$

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

$$\sum_{i=1}^{n} \sum_{k=1}^{K} p_{ik}\, log\, \pi_{ik} \quad (23)$$

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

In the current study, ML-EM estimation of CACE was carried out by the Mplus program (Muthén & Muthén, 1998). Parametric standard errors are computed from the information matrix of the ML estimator using both the first- and second-order derivatives under the assumption of normally distributed outcomes. For more details about estimation procedures in general latent variable modeling, see Muthén and Shedden (1999) and the chapter authored by Muthén in this book. Aslo, check Mplus website (www.statmodel.com) for more examples.

ESTIMATION OF CACE WITH A SINGLE CONTINUOUS OUTCOME

This section demonstrates the estimation of CACE with a single continuous outcome using the Job Search Intervention Study for unemployed workers (JOBS II; Vinokur, Price, & Schul, 1995; Vinokur & Schul, 1997). The JOBS II Intervention Study is a randomized field experiment intended to prevent poor mental health and promote high-quality reemployment. The experimental condition consisted of five half-day training seminars, which included the application of problem-solving and decision-making group processes, inoculation against setbacks, provision of social support and positive regard from the trainers, and learning and practicing job search skills. The control condition consisted of a booklet briefly describing job search methods and tips.

TABLE 3.1
JOBS II: Sample Statistics (N = 486)

Variable	*M*	*SD*	*Description*
TX	0.67	0.47	Experimental condition (0 = control, 1 = treatment)
c	0.55	0.50	Compliance (0 = noncompliance, 1 = compliance) in TX group
Depress0	2.45	0.30	Depression level before TX
Depress6	2.01	0.73	Depression level 6 months after TX
Employ6	0.62	0.49	Employment status 6 months after TX (0 = unemployed, 1 = employed)
Age	36.61	10.04	Age in years
Motivation	0.32	0.47	Motivation level before TX (0 = low, 1 = high)
Education	13.37	2.01	School grade completed
Assertive	3.07	0.91	Assertiveness before TX
Nonmarried	0.62	0.49	Marital status (0 = married, 1 = other)
Econ-Hard	3.60	0.87	Economic hardship before TX
Non-White	0.19	0.39	Race (0 = white, 1 = other)
Female	0.58	0.49	Gender (0 = male, 1 = female)

The present study focused on the high-risk status group based on previous studies (Price, van Ryn, & Vinokur, 1992; Vinokur, Price, & Shul, 1995), which indicated that the job search intervention had its primary impact on high-risk respondents. Risk score was computed based on risk variables predicting depressive symptoms at follow-up (depression, financial strain, and assertiveness) in the screening data (Price et al., 1992). A total sample size of 486 was analyzed in this study after listwise deletion of cases that had missingness in covariates and outcome variables. The variables used in the current study are described in Table 3.1.

Depression and reemployment are the major outcome measures in the JOBS II intervention study. The level of depression 6 months after the intervention (Depress6) is used as a continuous outcome measure in this section. The effect of the intervention on reemployment is analyzed in a later section. Depression was measured with a subscale of 11 items based on the Hopkins Symptom Checklist (Derogatis, Lipman, Rickles, Uhlenuth, & Covi, 1974).

Table 3.2 shows the results from the JOBS II data analysis using the ITT approach. In this method, it is assumed that noncompliers receive the same effects from the intervention as compliers. Table 3.2 shows that there is a small and insignificant effect of the intervention on the level of depression (TX effect = −0.137, Effect size = 0.189). The effect size of the treatment is calculated by dividing the outcome difference in treatment and control condition means by the square root of the variance pooled across the control and treatment groups. In the ITT analysis, economic hardship was found to be a significant predictor of the level of depression. Individuals had a higher level of depression if they had economic hardship.

TABLE 3.2
Intervention Effects on Depression: ITT Analysis

Parameter	*Estimate*	*SE*
Average treatment effects on Depress6	−0.137	0.072
Depress6 Regressed on x		
Depress0	0.063	0.108
Age	0.000	0.003
Motivation	0.019	0.073
Education	−0.026	0.016
Assertive	−0.039	0.038
Nonmarried	−0.117	0.075
Econ-Hard	0.143	0.040
Non-White	0.057	0.092
Female	0.105	0.068
Intercept	1.895	0.389
σ_{ε}^2	0.502	0.036

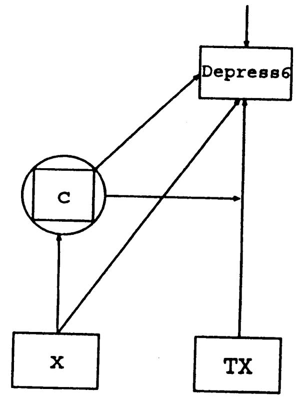

FIG. 3.2. CACE estimation with a single continuous outcome.

Figure 3.2 illustrates the model to estimate differential treatment effects in the JOBS II Intervention Study using the CACE approach. This model has been previously analyzed by Little & Yau (1998) using the ML-EM, treating unknown compliance status as missing data (Little & Rubin, 1997). In this method, compliance status of control group individuals is estimated, and average causal effects of the treatment are estimated only for compliers. In this diagram, TX denotes treatment assignment (0 = control, 1 = treatment) and c denotes compliance status (0 = noncompliance, 1 = compliance). Individuals who completed at least one seminar were categorized as *compliers* (55% of treatment group individuals) and the rest were categorized as *noncompliers*. Here the compliance status of the control group is latent (unknown), and the compliance status of the treatment group is observed (known). The partly observed latent variable c is expressed as a square in a circle. In the path diagram in Fig. 3.2 and in the other path diagrams to appear later, squares represent observed variables and circles represent latent (missing) variables. The path from TX to y corresponds to the treatment effect. The arrow from c to this path indicates that the treatment effect is different depending on compliance status. The arrow from c to y means that the means are different between compliers and noncompliers in the control group. In this model, covariates (x) including baseline depression (Depress0) are used as predictors of not only the outcome measure (Depress6) but also the compliance status (c) to improve precision in the prediction of compliance status and the quality of the treatment effect estimates.

Table 3.3 shows the results from the CACE analysis of the JOBS II intervention. In the current study, effect sizes of CACE estimates were calculated in a conventional way by dividing the outcome difference in treatment and control condition means by the square root of the variance pooled across the control and treatment groups. A more correct way to calculate effect size is to use the pooled variance of each compliance class. However, this approach was not chosen

TABLE 3.3
Intervention Effects on Depression: CACE Analysis

Parameter	*Estimate*	*SE*
CACE	−0.351	0.139
Depress6 Regressed on x		
Depress0	0.065	0.107
Age	−0.001	0.004
Motivation	−0.002	0.076
Education	−0.030	0.017
Assertive	−0.040	0.038
Nonmarried	−0.120	0.076
Econ-Hard	0.151	0.041
Non-White	0.065	0.092
Female	0.099	0.069
Intercept (Complier)	2.133	0.425
Intercept (Noncomplier)	1.821	0.385
σ_y^2	0.490	0.036
c Regressed on x (Complier vs. Noncomplier)		
Depress0	−0.420	0.425
Age	0.078	0.015
Motivation	1.309	0.292
Education	0.304	0.071
Assertive	−0.338	0.149
Nonmarried	0.546	0.288
Econ-Hard	−0.225	0.155
Non-White	−0.424	0.330
Female	−0.396	0.259
Intercept	−4.208	1.623

because standard deviations may vary depending on CACE models specified to estimate treatment effects, and this makes the comparison between models very difficult.

Table 3.3 shows that the intervention had a positive impact on the level of depression for compliers (TX effect = −0.351, Effect size = 0.484). In this method, the treatment effect is significant, and its magnitude is much larger than that of the overall average effects in the ITT analysis (e.g., Effect size = 0.189). The level of depression is significantly lower for compliers in the intervention condition compared with that of control condition individuals who could have complied if they had been assigned to the intervention condition. In the CACE analysis, economic hardship was found to be a significant predictor of the level of depression. It was also found that subjects complied more if they were older, more motivated, more educated, and less assertive.

The difference in the results from the ITT approach (Table 3.2) and those from the CACE approach (Table 3.3) implies that quite different conclusions are possible depending on the estimation method used to evaluate the effect of intervention treatment. According to the ITT analysis, the intervention did not have a significant effect on depression, and the magnitude of the effect was trivial. In contrast, the CACE analysis showed that the intervention had a significant effect on depression level for compliers and had a practically meaningful effect size.

ESTIMATION OF CACE WITH A SINGLE CATEGORICAL OUTCOME

This section demonstrates the estimation of CACE with a single categorical outcome using the Study of Vitamin Supplement Effect on Survival Rates in young children in Indonesia (Aceh Study; Imbens & Rubin, 1997; Sommer & Zeger, 1991; Sommer et al., 1986). The Aceh Study is a large-scale randomized controlled community trial conducted through a joint collaboration of the Dana Center for Preventive Ophthalmology at Johns Hopkins University, Hellen Keller International, and the Indonesian government in a province (Aceh) in Indonesia. The major goal of the Aceh Study is to examine the effectiveness of the intervention in reducing the mortality rate among infants and young children due to vitamin A deficiency. The study was originally aimed for children from 12 to 85 months old, but some children under 12 months or over 85 months old were also included in the study. Therefore, the effect of the age of children needs to interpreted with caution in this study. In the intervention condition villages, there were village-based persons trained by the government to give out the capsules. They were supposed to give each child a capsule every 6 months. Parents were asked at the end of 1 year of intervention whether their children had received a vitamin A capsule in the past 6 months. A total sample size of 20,130 was analyzed in this

TABLE 3.4
Aceh Study: Sample Statistics ($N = 20{,}130$)

Variable	*M*	*SD*	*Description*
TX	0.52	0.50	Experimental condition (0 = control, 1 = treatment)
c	0.81	0.39	Compliance (0 = noncompliance, 1 = compliance) in TX group
Survival	0.993	0.08	Vital status at 1 year follow-up (0 = died, 1 = alive)
Age	37.55	20.97	Age in months
Male	0.51	0.50	Gender (0 = female, 1 = male)
SES	0.63	0.48	Land ownership (0 = does not own land, 1 = owns land)
Health	0.87	0.33	Health seeking in the past year by any household member (0 = no, 1 = yes)
Diepast	0.64	1.07	Number of children whom the mother has had died in the past
Nblind	0.008	0.09	Nightblindness in past six months (0 = no, 1 = yes)

study after listwise deletion of cases that had missingness in covariates and outcome variables. The variables used in the current study are described in Table 3.4.

In the Aceh Study, the vital status of children at 1-year follow-up is the major outcome measure and is used as a binary outcome in this section. Vital status was measured at the end of 1 year of intervention. Sixty children died in the first 6 months of the trial, 75 children died in the second 6 months of the trial, and 19,995 children were alive at the end of the trial. Children who died either in the first or second 6-month trials were categorized as *not survived*, and children who were alive at the end of the trial were categorized as *survived* in this study. The survival rate among 10,439 intervention condition children was 0.995, and the survival rate among 9,691 control condition children was 0.992.

Table 3.5 shows the results from the Aceh Study data analysis using the ITT approach. In this method, it is assumed that noncompliers receive the same effects from the intervention as compliers. Table 3.5 shows that the intervention had a significant effect on survival rates of young children (TX effect = 0.446, Odds ratio = 1.561). The logistic regression results show that the odds of survival are 1.561 times higher for children in the intervention condition than for children in the control condition. In the ITT analysis, child's age and mortality rate of child's siblings were found to be significant predictors of the survival rate. Children had a higher rate of survival if they were older and had fewer siblings who had died.

Figure 3.3 illustrates the model to estimate differential treatment effects in the Aceh Study using the CACE approach. The CACE estimation of the intervention effects in the Aceh Study has been previously analyzed without covariates using EM and data augmentation algorithms (Imbens & Rubin, 1997). The current study employs the EM algorithm and incorporates covariates in the model. For CACE estimation of the intervention, a dichotomous variable (c) was created based on the dosage of vitamin A each child had taken. Children who took one or two capsules were categorized as *compliers* (81% of intervention condition children) and the rest were categorized as *noncompliers*. In this model, covariates (x) are used as predictors of not only the outcome measure (Survival) but also the

TABLE 3.5
Intervention Effects on Survival: ITT Analysis

Parameter	*Estimate*	*SE*
Average treatment effects on survival	0.446	0.177
Age	0.046	0.006
Male	−0.194	0.174
SES	0.064	0.177
Health	0.036	0.254
Diepast	−0.264	0.052
Nblind	−1.400	0.727
Intercept	4.933	0.290

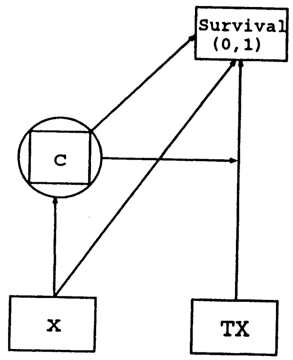

FIG. 3.3. CACE estimation with a single categorical outcome.

TABLE 3.6
Intervention Effects on Survival: CACE Analysis

Parameter	*Estimate*	*SE*
CACE	0.813	0.274
Survival Regressed on x		
Age	0.043	0.005
Male	−0.189	0.174
SES	−0.014	0.182
Health	0.015	0.253
Diepast	−0.261	0.052
Nblind	−1.436	0.734
Intercept (Complier)	4.008	0.346
Intercept (Noncomplier)	3.351	0.342
c Regressed on x (Complier vs. Noncomplier)		
Age	0.005	0.001
Male	0.000	0.050
SES	0.394	0.051
Health	0.097	0.077
Diepast	0.004	0.025
Nblind	0.223	0.313
Intercept	0.938	0.094

compliance status (c) to improve precision in the prediction of compliance status and the quality of the treatment effect estimates.

Table 3.6 shows the results from the Aceh Study data analysis using the CACE approach. The logistic regression of vital status in the CACE approach shows that the intervention had a significant effect on survival rates of young children (TX effect = 0.813, Odds ratio = 2.254), and the odds ratio is considerably higher than in the ITT approach (i.e., 1.561). The odds of survival are 2.254 times

higher for intervention condition children who actually took capsules than for control condition children who could have taken capsules if they had been assigned to the intervention condition. In the CACE analysis, child's age and mortality rate of child's siblings were found to be significant predictors of the survival rate. Children had a higher rate of survival if they were older and had fewer siblings who had died. It was also found that parents complied with the intervention more if they had higher socioeconomic status (SES) and if their children were older.

In the Aceh Study, both ITT and CACE approaches showed significant effects of the intervention on the vital status of young children. However, the magnitudes of the intervention effects are quite different in two approaches. These results imply that treatment effect estimates for categorical outcomes could be still sensitive to estimation method in the presence of noncompliance, although noncompliance rate is quite low (19%) and the sample size is very large.

ESTIMATION OF CACE WITH MULTIPLE OUTCOMES

This section demonstrates the estimation of CACE with multiple outcome measures using the JOBS II Study. The same subset of the JOBS II data with a sample size of 486 used earlier was analyzed in this section. The variables used in this section are described in Table 3.1. This section focuses on the estimation of the intervention effects on reemployment, which was one of the major goals of the JOBS II intervention. Reemployment status was determined 6 months after the intervention by classifying respondents working for 20 hours or more per week as reemployed (Employ6 = 1) and the rest as unemployed (Employ6 = 0).

Table 3.7 shows the results from the CACE analysis using a single categorical outcome (Employ6). The logistic regression of reemployment status in the CACE approach shows that the intervention did not have a significant effect on reemployment among intervention condition individuals, although they actually had complied with the intervention (TX effect = 0.576, Odds ratio = 1.779). In the CACE analysis using a single categorical outcome, it was found that age, education, and racial background were significant predictors of the reemployment. Individuals had a higher rate of reemployment if they were White, younger, and more educated. It was also found that individuals complied more if they were older, single, more motivated, more educated, and less assertive.

Figure 3.4 illustrates the model to estimate CACE using multiple outcomes in the JOBS II Intervention Study. In this method, compliance status (c) of control group individuals is estimated based on both outcomes, and intervention effects for compliers are also estimated for both outcomes. The binary and continuous outcomes are correlated through covariates, intervention assignment, and compliance status, but there is no direct relation between the binary outcome and the residual of the continuous outcome. The conditional independence between these

TABLE 3.7
Intervention Effects on Employment: CACE Analysis

Parameter	*Estimate*	*SE*
CACE	0.576	0.344
Employ6 Regressed on x		
Depress0	−0.058	0.323
Age	−0.023	0.011
Motivation	−0.267	0.215
Education	0.130	0.053
Assertive	0.075	0.118
Nonmarried	0.201	0.210
Econ-Hard	0.015	0.122
Non-White	−0.554	0.253
Female	−0.051	0.202
Intercept (Complier)	−0.761	1.290
Intercept (Noncomplier)	−0.499	1.191
c Regressed on x (Complier vs. Noncomplier)		
Depress0	−0.387	0.428
Age	0.078	0.016
Motivation	1.244	0.294
Education	0.301	0.070
Assertive	−0.347	0.152
Nonmarried	0.571	0.291
Econ-Hard	−0.239	0.157
Non-White	−0.386	0.331
Female	−0.386	0.260
Intercept	−4.193	1.648

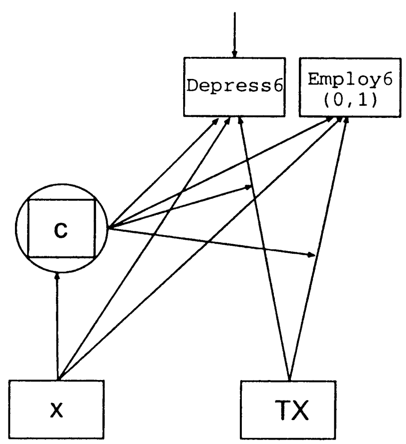

FIG. 3.4. CACE estimation with multiple outcomes.

TABLE 3.8
Intervention Effects on Employment and Depression: CACE Analysis

Parameter	*Estimate*	*SE*	*Estimate*	*SE*
	Employ6		*Depress6*	
CACE	0.693	0.353	−0.369	0.142
	Employ6 Regressed on x		*Depress6 Regressed on x*	
Depress0	−0.062	0.324	0.066	0.107
Age	−0.022	0.011	−0.001	0.004
Motivation	−0.237	0.219	−0.005	0.077
Education	0.137	0.053	−0.031	0.017
Assertive	0.070	0.118	−0.038	0.038
Nonmarried	0.205	0.211	−0.119	0.075
Econ-Hard	0.007	0.123	0.152	0.041
Non-White	−0.565	0.253	0.064	0.092
Female	−0.053	0.203	0.100	0.069
Intercept (Complier)	−0.980	1.289	2.161	0.423
Intercept (Noncomplier)	−0.532	1.191	1.820	0.385
σ_y^2			0.488	0.036
	c Regressed on x (Complier vs. Noncomplier)			
	Estimate	*SE*		
Depress0	−0.450	0.422		
Age	0.078	0.015		
Motivation	1.303	0.291		
Education	0.306	0.070		
Assertive	−0.351	0.151		
Nonmarried	0.521	0.289		
Econ-Hard	−0.223	0.156		
Non-White	−0.384	0.333		
Female	−0.403	0.257		
Intercept	−4.119	1.616		

two outcome measures is assumed for the simplicity in the model estimation, but this assumption may need to be relaxed. In this model, two major outcomes (Employ6 and Depress6) of the intervention are considered at the same time to improve the quality of parameter estimates in the categorical outcome (Employ6). The model is intended to increase the precision in the estimation of compliance status in the control condition by including a continuous outcome (Depress6), and consequently to increase the power to detect intervention effects on the categorical outcome (Employ6).

Table 3.8 shows the results from the CACE analysis of the JOBS II intervention using multiple outcomes illustrated in Fig. 3.4. The logistic regression of reemployment status shows that the intervention had a positive effect on reemployment for compliers (TX effect = 0.693, Odds ratio = 2.000). In this method, the intervention effect is significant, and its magnitude is larger than that in the CACE analysis using a single categorical outcome only. The odds of reemployment are two times higher for intervention condition individuals who actually participated in intervention seminars than for control condition individuals who could have participated if they had been assigned to the intervention condition. The logistic regression of reemployment status also shows that age, education, and racial background were significant predictors of the reemployment. Individuals had a higher rate of reemployment if they were White, younger, and more educated.

Table 3.8 also shows the estimation of intervention effects on the level of depression 6 months after the intervention. The results show that the intervention had a positive effect on depression for compliers (TX effect = −0.369, Effect size = 0.509). The intervention effects on depression are slightly stronger in this model than in the CACE model using a continuous outcome only (see Table 3.3). Among several covariates, economic hardship was found to be a significant predictor of the level of depression. It was also found that subjects complied more if they were older, more motivated, more educated, and less assertive.

The difference in the results from the CACE approach using a single outcome (Tables 3.3 and 3.7) and those from the CACE approach with multiple outcomes (Table 3.8) implies that the efficiency in CACE estimation can be improved by employing estimation models based on multiple outcomes. The difference between the two methods was not dramatic, but still affected the power to detect intervention effects.

GROWTH MIXTURE CACE ANALYSIS FOR MULTIPLE OUTCOMES WITH A TREND

This section demonstrates CACE estimation using repeated outcome measures with a trend using the Johns Hopkins Public School Preventive Intervention Study. In the previous section, multiple outcome measures are used in CACE estimation, but these outcomes were not repeated measures of the same outcome. When intervention studies are focused on the long-term effects of treatment, the outcome is often measured several times at specific intervals. In this case, one way to define the treatment effect is to use the difference between the treatment and the control group in the outcome measured at the last time point, conditioning on the outcome measured at the first time point (ANCOVA). Another way to define the treatment effect is to use a trend or growth trajectory of the subjects. This section demonstrates CACE estimation in these two alternative approaches.

The Johns Hopkins Public School Preventive Intervention Study was conducted by the Johns Hopkins University Preventive Intervention Research Center in 1993–1994 (Ialongo et al., 1999). The study was designed to improve academic achievement and reduce early behavioral problems of school children. Based on

the life course/social field framework as described by Kellam and Rebok (1992), the study focused on successful adaptation to first grade as a means of improving social adaptational status over the life course. Teachers and first-grade children were randomly assigned to intervention conditions. The intervention impact was assessed in the spring of first and second grades. Two intervention programs were employed in the Johns Hopkins Public School Preventive Intervention Study: the Classroom-Centered Intervention and the Family-School Partnership Intervention. The present study focused on the comparison between the control group and the Family-School Partnership Intervention group. Intervention condition parents were asked to implement 66 take-home activities related to literacy and mathematics. Based on the level of completeness in home-learning activities, a dichotomous variable was created in this study. Parents who completed at least 35 activities were categorized as *compliers* (73% of parents) and the rest were categorized as *noncompliers*. The cutpoint was decided based on exploratory growth mixture analyses (Jo & Muthén, 2000), but the details are not discussed in this chapter. For illustration purpose, compliance in continuous measure was simply dichotomized in this example, but note that sensitivity of the CACE estimate to different thresholds needs to be carefully examined in practice (West & Sagarine, 2000). Figure 3.5 shows observed mean curves of attention deficit in the Johns Hopkins School Preventive Intervention Study.

A total sample size of 286 was analyzed in this study after listwise deletion of cases that had missingness in covariates and outcome variables. The two major outcome measures in the Johns Hopkins Public School Preventive Intervention Study were academic achievement (CTBS mathematics and reading test scores) and the score Teacher Observation of Classroom Adaptation–Revised (TOCA–R) score (Werthamer-Larsson, Kellam, & Wheeler, 1991). Among these two outcome meas-

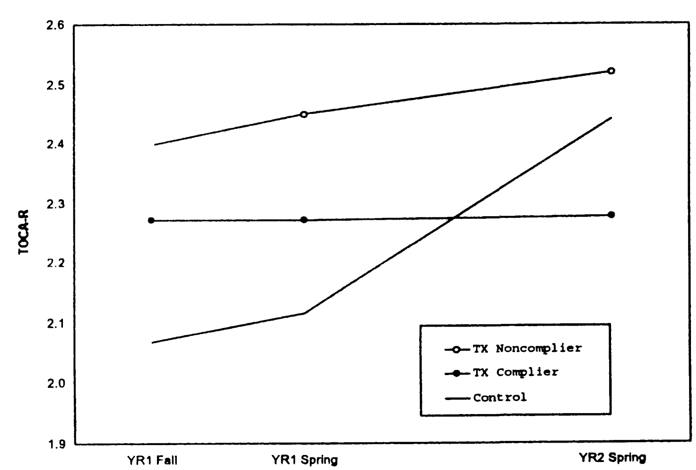

FIG. 3.5. Observed mean curves of attention deficit.

TABLE 3.9
The Johns Hopkins Public School Prevention Data: Sample Statistics ($N = 286$)

Variable	*M*	*SD*	*Description*
TX	0.50	0.50	Experimental condition (0 = control, 1 = treatment)
Activity	40.78	15.96	Number of completed home-learning activities
c	0.73	0.45	Dichotomized home-learning activities (0 = 34 or fewer, 1 = 35 to 66)
AD0	2.19	0.92	TOCA teacher report mean attention deficit before TX (1st grade, fall)
AD6	2.22	0.95	TOCA teacher report mean attention deficit 6 month after TX
AD18	2.39	1.11	TOCA teacher report mean attention deficit 18 month after TX
Male	0.49	0.50	Student's gender (0 = female, 1 = male)
Lunch	0.62	0.49	Free lunch program (0 = no, 1 = yes)
Page	3.01	1.44	Parent's age in 5-year brackets
Pmale	0.07	0.26	Parent's gender (0 = female, 1 = male)
Non-White	0.87	1.03	Parent's ethnicity (1 = non-White, 0 = White)

TABLE 3.10
Intervention Effects on Attention Deficit: CACE Analysis

Parameter	*Estimate*	*SE*
CACE	−0.300	0.129
AD18 Regressed on x		
AD0	0.624	0.068
Male	0.330	0.107
Lunch	0.296	0.112
Page	0.037	0.037
Pmale	0.132	0.258
Non-White	−0.213	0.146
Intercept (Complier)	0.890	0.246
Intercept (Noncomplier)	0.765	0.276
σ_y^2	0.790	0.074
c Regressed on x (Complier vs. Noncomplier)		
AD0	−0.118	0.226
Male	0.203	0.390
Lunch	−0.147	0.382
Page	−0.197	0.126
Pmale	−0.781	0.637
Non-White	−0.306	0.639
Intercept	2.209	0.782

ures, the TOCA–R score was used as the final outcome measure in this study. 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 several areas that TOCA–R measures, attention deficit is the construct focused in this study. The attention deficit scale ranges from 1 to 6 and consists of TOCA–R items that measure hyperactivity, concentration problems, and impulsiveness. Table 3.9 shows the sample statistics for the variables used in the analyses of this study.

Table 3.10 shows the results from the CACE analysis using a single outcome measured approximately 18 months after the intervention (AD18). In this approach, the outcome measured before the treatment (AD0) is used as one of covariates, and the outcome measured in the spring of the first grade (AD6) is ignored (i.e., ANCOVA). The results show that the intervention had a positive impact on children's attention deficit when their parents were highly involved in the intervention activities (TX effect = −0.300, Effect size = 0.271). It was assumed that there was no effect of intervention assignment for children with parents who had a very low level of compliance with the intervention activities, but this assumption may need to be relaxed. The assumption of the exclusion restriction is critical for the identifiability of CACE models, but can be unrealistic in some situations (Hirano et al., 2000; Jo, 2000a, 2000b). In the CACE analysis based on a single outcome measure, baseline attention deficit, gender, and free lunch program were found to be significant predictors of the level of attention deficit. Children had a higher level of attention deficit in spring of the second grade if their baseline attention deficit was higher, if they were boys, and if their SES level was low.

Figure 3.6 illustrates the growth mixture CACE model using repeated outcome measures. In this approach, all three measures of attention deficit are considered

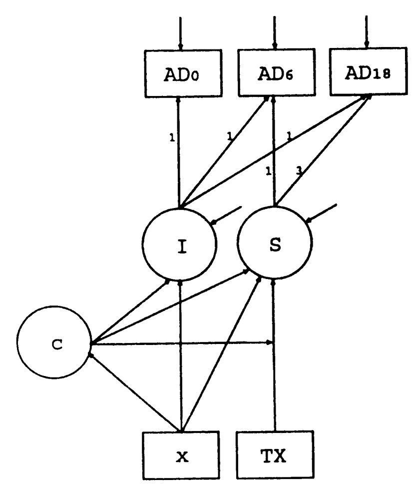

FIG. 3.6. Growth mixture CACE estimation with repeated outcome measures.

in the analysis. This approach is in line with the CACE approach using a single outcome measure in the sense that compliance status of control group individuals is estimated, and average causal effects of the treatment are estimated only for compliers. One difference between CACE models using growth mixtures and CACE models using a single outcome is that the first time point measure (AD0) is one of the outcome measures instead of one of the covariates. Because initial status (I) and growth rate (S) are separated in this model, the influence of background variables can be estimated separately for initial level of attention deficit and change of attention deficit. Another difference is that the growth mixture CACE model utilizes not only covariates, but also trajectory information to identify compliance class and increase efficiency in the estimation of intervention effects. Including a growth process in the estimation of CACE utilizes the idea of

TABLE 3.11
Intervention Effects on Attention Deficit: Growth Mixture CACE Analysis

Parameter	*Estimate*	*SE*
CACE	−0.306	0.126
Growth rate Regressed on x		
Male	0.062	0.038
Lunch	0.063	0.039
Page	0.013	0.013
Pmale	0.061	0.084
Non-White	−0.107	0.050
Intercept (Complier)	0.103	0.074
Intercept (Noncomplier)	0.047	0.087
Initial Status Regressed on x		
Male	0.376	0.102
Lunch	0.277	0.104
Page	−0.002	0.036
Pmale	−0.131	0.128
Non-White	0.268	0.159
Intercept (Complier)	1.569	0.207
Intercept (Noncomplier)	1.673	0.234
ψ_S	0.042	0.011
ψ_I	0.536	0.061
σ_v^2	0.265	0.030
c Regressed on x (Complier vs. Noncomplier)		
Male	0.144	0.387
Lunch	−0.176	0.385
Page	−0.198	0.127
Pmale	−0.775	0.632
Non-White	−0.327	0.625
Intercept	2.025	0.713

a general latent variable modeling framework, where both categorical and continuous latent variables are incorporated (Muthén, 1998; Muthén et al., 1997; Muthén & Shedden, 1999). 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.

In Fig. 3.6, initial status I has equal loadings $(1, 1, 1)$ on three outcome measures representing initial status, which does not change over time. The time scores (h) are fixed at 0, 1, and 3 representing linear growth over time. The arrows from c to I and S mean that trajectory shapes are different between compliers and noncompliers in the control group. The arrow from TX to S corresponds to the mean shift in growth rate due to the treatment. The arrow from c to this path indicates that the treatment effect is different depending on the compliance status. The intervention effects for compliers (CACE) is defined as the difference in estimated attention deficit between the control and the treatment condition at the last time point (see Eq. [16]).

Table 3.11 shows the results from the estimation of treatment effects using growth mixture CACE modeling. The results show that the intervention had a positive impact on children's attention deficit change when their parents were highly involved in the intervention activities (TX effect = 0.306, Effect size = 0.276). It is also shown that growth mixture CACE analysis has a slightly larger effect size and tighter confidence interval than the CACE analysis using the ANCOVA approach shown in Table 3.10. In the growth mixture CACE analysis, child's gender and participation in the free lunch program were significant predictors of initial level of attention deficit, and parents' racial background was a

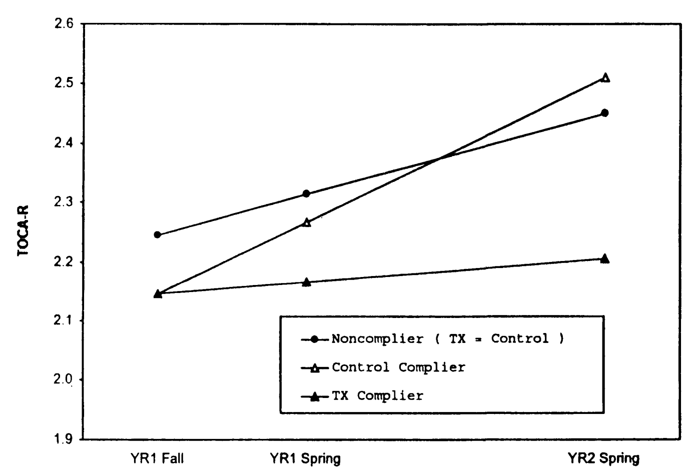

FIG. 3.7. Estimated mean curves of attention deficit.

significant predictor of growth rate of attention deficit. Initial level of attention deficit was higher for boys with low SES. The level of attention deficit increased significantly faster for children from White families.

Figure 3.7 shows estimated mean attention deficit curves over time based on results in Table 3.11. This figure shows how attention deficit changed over time depending on parents' compliance level and treatment assignment. It is shown that attention deficit among highly complying parents' children maintained a low level over time, but the deficit could increase to a level even higher than that of less involved parents' children at the second grade unless the intervention was given.

CONCLUSION

Noncompliance is a common problem in intervention studies, and one can arrive at different conclusions about the effect of the same intervention trial depending on how this problem is handled. Both ITT and CACE analyses are useful in their own contexts. However, the estimation of CACE was the focus of this study, because a major interest in intervention trials is often the estimation of treatment effects for individuals who actually receive the treatment.

The current study demonstrated that the problem of noncompliance can be dealt with in a broader framework of structural equation modeling by looking at compliance status as a categorical latent variable. To deal with compliance status as a latent variable, a broader framework of structural equation modeling was employed. This framework has two differences in the concept of latent variable from the conventional structural equation modeling. First, latent variable can be not only continuous but also categorical, whereas latent variable is only continuous in the conventional framework. Second, *latent* may mean missing for only a part of the total sample, whereas it usually means missing or unknown for everybody in the conventional framework.

This study demonstrated that the general latent variable approach is useful in improving the efficiency and interpretability of CACE estimation. Possibilities of flexible CACE modeling in a general latent variable modeling framework were demonstrated in various situations. The examples shown in this study imply that the difference in the estimation of treatment effects could be substantial depending not only on estimation approaches, but also modeling alternatives.

In the examples of intervention effect estimation using a single outcome measure, it was shown that the magnitude of treatment effects was considerably larger in the CACE approach than in the ITT approach. In the JOBS II study example using a continuous outcome measure, the intervention did not have a significant effect on depression, and the magnitude of the effect was trivial according to the ITT analysis. In contrast, the CACE analysis showed that the intervention had a significant effect on depression level for compliers and had a practically meaningful effect size. In the Aceh Study example using a

categorical outcome measure, both ITT and CACE approaches showed significant effects of the intervention on vital status of young children. However, the magnitudes of the intervention effects were quite different in two approaches, implying that treatment effect estimates for categorical outcomes could still be sensitive to estimation method in the presence of noncompliance even with a high compliance rate and a large sample size. In both examples, covariates were incorporated in CACE models to increase the precision in the estimation of compliance class and improve the power to detect treatment effects.

This study also demonstrated the use of multiple outcomes and growth trajectories in the estimation of CACE. It was shown that the quality of intervention effect estimates could be improved further within the CACE approach by employing models that utilize the information from multiple outcomes and growth trajectories. In the CACE estimation of the JOBS II study, the intervention effect on reemployment status was not significant when reemployment status was the only outcome in the model. In contrast, the intervention effect was significant and its magnitude was larger when both outcomes (reemployment and depression) were included in the model. In the CACE estimation of the Johns Hopkins study, the difference between ANCOVA and growth mixture approaches was small in terms of the magnitude of the intervention effects. However, the growth mixture CACE approach provided more detailed information about the intervention effects. It was found that the intervention had a positive impact on attention deficit among highly complying parents' children. Initial level of attention deficit was higher for boys with low SES. The level of attention deficit increased significantly faster for children from White families.

APPENDIX

Mplus input for Table 3.3

```
title: CACE estimation with a single continuous outcome
data: file is jobs2.dat;
variable: names are depress6 TX depress0 age motivat educat;
          names are assert nonmarr econhard nonwhite female employ6 c1 c2;
          usev are depress6 TX depress0-female c1 c2;
          classes-c(2);
          training=c1-c2;
analysis: type=mixture;
model:
%OVERALL%
          C#1 ON depress0-female;
          depress6 ON TX depress0-female;
          depress6;
          [depress6];
```

```
%C#1%
        [depress6];
        depress6 ON TX;
%C#2%
        [depress6];
        depress6 ON TX@0;
```

Mplus input for Table 3.6

```
title: CACE estimation with a single categorical outcome
data: file is aceh.dat;
variable: names are survival TX age male ses health diepast nblind c1 c2;
        categorical are survival;
        classes = c(2);
        training = c1-c2;
analysis: type = mixture;
model:
%OVERALL%
        C#1 ON age-nblind;
        survival ON TX age-nblind;
%C#1%
        [survival$1*-4];
        survival ON TX;
%C#2%
        [survival$1*-3];
        survival ON TX@0;
```

Mplus input for Table 3.8

```
title: CACE estimation with multiple outcomes
data: file is jobs2.dat;
variable: names are depress6 employ6 TX depress0-female c1 c2;
        categorical are employ6;
        classes = c(2);
        training = c1-c2;
analysis: type = mixture;
model:
%OVERALL%
        C#1 ON depress0-female;
        depress6 ON TX depress0-female;
        employ6 ON TX depress0-female;
        [depress6];
        depress6;
```



```
%C#1%
        [depress6];
        [employ6$1*1.0];
        depress6 ON TX;
        employ6 ON TX;
%C#2%
        [depress6];
        [employ6$1*0.5];
        depress6 ON TX@0;
        employ6 ON TX@0;
```

Mplus input for Table 3.11

```
title: CACE estimation with repeated outcome measures
data: file is hopkins.dat;
variable: names are TX AD0 AD6 AD18;
        names are male lunch page pmale nonwhite c1 c2;
        usev are AD0 AD6 AD18 TX male-nonwhite c1 c2;
        classes = c(2);
        training = c1-c2;
analysis: type = mixture;
model:
%OVERALL%
        init by AD0-AD18@1;
        grow by AD0@0 AD6@1 AD18@3;
        [AD0-AD18@0];
        AD0-AD18 (1);
        grow ON TX male-nonwhite;
        init ON male-nonwhite;
        init;
        [init];
        grow;
        [grow];
        C#1 ON male-nonwhite;
%C#1%
        [init];
        [grow];
        grow ON TX;
%C#2%
        [init];
        [grow];
        grow ON TX@0;
```

ACKNOWLEDGEMENT

This study is supported by the NIMH grant P30MH38330, Michigan Prevention Research Center at the Institute for Social Research, Richard H. Price, P.I., University of Michigan. This study is also supported by grant 1 KO2 AA 00230-01 from NIAAA and by NIMH grant P50 MH38725, Epidemiologic Prevention Center for Early Risk Behaviors, Philip Leaf, P.I., Department of Mental Hygiene, Johns Hopkins School of Hygiene and Public Health. We would like to thank Amiram Vinkour, Nicholas Ialongo, and Joanne Katz for their data and helpful comments.

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 non-compliers 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* (Series B), *39*, 1–38.

Derogatis, L. R., Lipman, R. S., Rickles, K., Uhlenuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL). In P. Pichot (Ed.), *Psychological measurements in psychopharmacology: Modern problems in pharmacopsychiatry* 7 (pp. 79–110). New York: Karger, Basel.

Hirano, K., Imbens, G. W., Rubin, D. B., & Zhou, X. H. (1999). *Assessing the effect of an influenza vaccine in an encouragement design*. Unpublished manuscripts.

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. *Sociological Methodology*, 449–484.

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. (2000a). *Estimating intervention effects with noncompliance: Alternative model specifications*. Submitted for publication.

Jo, B. (2000b). *Model misspecification sensitivity analysis in estimating causal effects of inverventions with noncompliance*. Submitted for publication.

Jo, B. (2000c). *Statistical power in randomized intervention studies wiht noncompliance*. Submitted for publication.

Jo, B., & Muthén, B. O. (2000). Intervention studies with noncompliance: Complier average causal effect estimation in growth mixture modeling. In N. Duan & S. Reise (Eds.), *Multilevel modeling: Methodological advances, issues, and applications, multivariate applications book series*. Mahwah, NJ: 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. Trembley (Eds.), Preventing anti-social behavior: interventions from birth through adolescence. New York; Guilford Press, pp. 162–195.

Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: John 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.

Muthén, B. O. (1998). 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. 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*. Unpublished manuscript.

Muthén, B. O., & Curran, P. J. (1997). General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation. *Psychological Methods, 3*, 371–402.

Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide*. Los Angeles: Author.

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

Price, R.H., van Ryn, M., Vinokur, A.D. (1992). Impact of a preventive job search intervention on the likelihood of depression among the unemployed. *Journal of Health and Social Behavior, 33*, 158–167.

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.

Sommer, A., Tarwotjo, I., Djunaedi E., et al. (1986). Impact of vitamin A supplementation on childhood mortality: A randomised controlled community trial. *Lancet, i*, 1169–1173.

Sommer, A., & Zeger, S. (1991). On estimating efficacy from clinical trials. *Statistics in Medicine, 10*, 45–52.

Tanner, M. (1996). *Tools for statistical inference: Methods for the exploration of posterior distributons 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: John Wiley.

Vinokur, A. D., Price, R. H., & Schul, Y. (1995). Impact of the JOBS intervention on unemployed workers varying in risk for depression. *American Journal of Community Psychology, 23*, 39–74.

Vinokur, A. D., & Schul, Y. (1997). Mastery and inoculation against setbacks as active ingredients in intervention for the unemployed. *Journal of Consulting and Clinical Psychology, 65*, 867–877.

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. Thousnad Oaks, CA: Sage. pp. 117–154.