Psychological Methods
1998, Vol. 3, No. 2, 147–159

1082-989X/98/$3.00

Statistical Techniques for Analyzing Data From Prevention Trials: Treatment of No-Shows Using Rubin's Causal Model

Roderick J. Little and Linda H. Y. Yau
University of Michigan

Noncompliance is a common challenge in the analysis and interpretation of prevention trials. The authors describe new formulations of the problem based on D. B. Rubin's (1974, 1978) causal model. The formulations help clarify assumptions underlying estimation procedures and yield more efficient methods of estimation. The authors apply the methods to a trial of a job training intervention in which nearly half the participants randomly assigned to the intervention failed to attend the job training seminars. An interesting feature is the presence of covariates measured prior to treatment randomization. Versions of the model that condition on these covariates suggest positive results for the intervention in a high-risk group but no evidence of gains in a low-risk group.

Prevention trials that assess the efficacy of interventions apply the interventions to groups of participants and then compare the distributions of relevant outcomes across groups. If statistically and substantively significant differences are found, the central issue is whether these differences can be attributed to causal effects of the interventions rather than to confounding factors. Randomized treatment assignment is a key design tool for limiting the undermining effects of confounding factors. The potential for bias in the assignment process is removed, and randomization balances the distribution of confounding factors across groups on average. Chance imbalances can be addressed by recording important covariates and controlling for them in the statistical analysis.

A number of practical problems undermine randomization and complicate the causal interpretation of observed treatment differences. Randomized assignment is not to be confused with quasi-random haphazard assignment and needs to be carefully conducted to ensure its integrity. When participants randomly assigned to a treatment fail to complete the study, completers are no longer random samples of the original groups. In this article we assume that randomization is carefully conducted and that data are fully recorded for all participants, and we focus on the problem of noncompliance, where participants fail to comply with the assigned treatment.

Consider a study in which participants are randomly assigned to a behavioral intervention for the prevention of AIDS or a control (null) treatment. Some fraction of those assigned to the intervention ignore the behavioral information and hence effectively receive the same information as controls. Other participants randomly assigned to the control group obtain the information that comprises the behavioral intervention from friends in the intervention group. An intent-to-treat analysis compares the distribution of outcomes between treatments as randomized, ignoring these lapses in compliance. A simple estimate of the average intent-to-treat effect is the difference in mean outcomes between those assigned to the treatment group and those assigned to the control group.

Roderick J. Little and Linda H. Y. Yau, Department of Biostatistics, University of Michigan. Linda H. Y. Yau is now at the Merck Research Laboratories, Blue Bell, Pennsylvania.

This work was conducted while Linda H. Y. Yau was a graduate student at the University of Michigan, and the work was supported by National Institute of Mental Health Grant P30 MH38330-12. While taking full responsibility for errors, we thank Rick Price and Amiram Vinokur for valuable support, encouragement, and suggestions, and Don Rubin for helpful comments.

Correspondence concerning this article should be addressed to Roderick J. Little, Department of Biostatistics, University of Michigan, 1420 Washington Heights, Ann Arbor, Michigan 48109-2029. Electronic mail may be sent to rlittle@umich.edu.

This estimates the causal effect of treatment assignment rather than the effect of the treatment for participants who actually received it. The estimator is protected from bias by randomized treatment allocation but as a measure of the effect of the intervention is distorted by the switching between groups.

Another approach is to classify participants according to the treatments actually received (an "as treated" analysis). The problem then is that randomization is violated, and confounding factors associated with switching potentially corrupt the causal interpretation of treatment effects. Because both of these approaches to analysis have problems, in practice both "as randomized" and "as treated" analyses are often carried out when the effects of treatments are being assessed.

We describe and apply here some recent work in statistics and econometrics (Angrist, Imbens, & Rubin 1996; Imbens & Rubin 1997a, 1997b) that provides a useful causal framework for thinking about noncompliance problems, as well as new and improved methods of statistical analysis. The focus is on another outcome, the average treatment effect for compliers, which Imbens and Rubin (1997b) called the *complier-average causal effect.* Because the degree of compliance is determined by the study participants, the extensive econometric literature on self-selection is also pertinent (e.g., Gronau, 1974; Heckman & Robb, 1985; Robins, 1989). In that approach, compliance status is included as a variable in a multivariate model, and instrumental variable methods of estimation are applied (Bowden & Turkington, 1984). We discuss here a simple version of that approach that yields an instrumental variable estimator of the complier-average causal effect known as *Bloom's method* by some prevention researchers (Bloom, 1984). This estimator also arises in the biostatistics literature on compliance (Sommer & Zeger, 1991).

We apply these approaches for estimating the complier-average causal effect to data from the JOBS II intervention trial (Price & Vinokur, 1995; Vinokur, Price, & Schul, 1995), which tested the efficacy of a job training intervention for unemployed people. The intervention consisted of five half-day job search seminars that helped the participants enhance their job search strategies. The control treatment was a self-guided booklet with advice on how to find a job. Noncompliance arose because nearly half of the participants randomized to the intervention did not attend the seminars. Vinokur, Price, and Caplan (1991) estimated the complier-average causal effect for the earlier JOBS I employment intervention trial using Bloom's (1984) estimator. In this article we apply that method and Imbens and Rubin's (1997a) approach to the JOBS II data set, with and without covariates.

Noncompliance as a Missing Data Problem

Consider a randomized study consisting of a new treatment and a control treatment. Imbens and Rubin (1997b) characterized study participants as compliers, always-takers, never-takers, and defiers. Compliers do what they are assigned to do, always-takers always take the new treatment regardless of assignment, never-takers never take the new treatment regardless of assignment, and defiers do the opposite of what they are assigned.

Let R_i be the randomization indicator for participant i, taking the value 1 if the ith participant is randomized to the new treatment and 0 if randomized to the control treatment. Let $T_i = T_i(R_i)$ be the indicator for the treatment that would actually be received by the ith participant when assigned R_i, taking the value 1 for treatment and 0 for control; the notation $T_i(R_i)$ reflects the dependence of the treatment received by participant i on the treatment assigned. Imbens and Rubin (1997b) actually adopted a more general notation that also allows the treatment received by participant i to depend on treatments assigned to other participants. For simplicity we exclude this possibility in our formulation. Thus:

if participant i is a complier, then $T_i(1) = 1$ and $T_i(0) = 0$;
if participant i is an always-taker, then $T_i(1) = T_i(0) = 1$;
if participant i is a never-taker, then $T_i(1) = T_i(0) = 0$;
if participant i is a defier, then $T_i(1) = 0$ and $T_i(0) = 1$.

In practice, knowledge of the compliance status of participants is incomplete, and Imbens and Rubin (1997b) treated compliance as a missing-data problem. Specifically, if a participant is assigned to the new treatment and complies, then $T_i(1) = 1$, and that participant may be a complier (if $T_i[0] = 0$) or an always-taker (if $T_i[0] = 1$). If a participant is assigned to new treatment and fails to comply, then $T_i(1) = 0$, and that participant may be a never-taker (if $T_i[0] = 0$) or a defier (if $T_i[0] = 1$). If a participant is assigned to the control condition and complies, then $T_i(0) = 0$, and that participant may be a complier (if $T_i[1] = 1$) or a never-taker (if $T_i[1] = 0$). If a participant is assigned to the control condition and obtains the new treatment, then $T_i(0) = 1$, and that participant may be

an always-taker (if $T_i[1] = 1$) or a defier (if $T_i[1] = 0$).

Let Y be an outcome of interest, and let $Y_i(R_i, T_i)$ be the potential outcome for a participant randomized to group R_i and receiving treatment T_i. This notation implies the *stable unit treatment value assumption* (SUTVA; Imbens & Rubin, 1997b; Rubin, 1978), defined as follows:

Assumption 1 (SUTVA): The potential outcomes for each individual i do not depend on the treatment status of other individuals in the sample.

An important conceptual feature of the framework is that causal effects are defined for individuals, even though individual effects cannot be directly observed. The two outcomes for participant i that are potentially observable are $Y_i[1, T_i(1)]$, if the participant is assigned to the new treatment, and $Y_i[0, T_i(0)]$, if the participant is assigned to the control group. The causal effect of treatment assignment R on Y for participant i is defined as the difference in these two quantities, that is, $Y_i[1, T_i(1)] - Y_i[0, T_i(0)]$ (Holland, 1986; Rubin, 1974, 1978). Note the following:

1. The causal effect of *treatment assignment* is not the causal effect of *treatment* unless there is full compliance, because only compliers always receive the treatments actually assigned.

2. The effect is generally not observable for individuals, because participants are assigned to intervention or control but not to both.

3. We can estimate average values of the causal effect of treatment assignment over groups of participants. Specifically, Let μ_1 be the population mean potential outcome when all participants are assigned to the new treatment and μ_0 be the population mean potential outcome when all participants are assigned to the control treatment. The population average causal effect of treatment assignment is then defined as $\delta = \mu_1 - \mu_0$. Let $\bar{y}_0$ and $\bar{y}_1$ denote the sample mean outcomes for participants randomly assigned to control and new treatment, respectively. Because of the random treatment assignment, $\bar{y}_1 - \bar{y}_0$ is an unbiased estimator of δ. This is an intent-to-treat analysis of the data.

We focus here on a different causal effect of interest, the complier-average causal effect, defined for a population as the average causal effect restricted to compliers ($C = 1$); Specifically,

$$\delta_c = \mu_{c1} - \mu_{c0},$$

where μ_{c1} and μ_{c0} are the mean outcomes when the new treatment and control respectively are applied to the population of compliers. If we knew the compliance status of all the participants in the study, this quantity could be estimated simply as the difference in new treatment and control means in the subsample of compliers, but the problem is that compliance status is unknown. Note that $\delta = \pi_c\delta_c + (1 - \pi_c)\delta_{\bar{c}}$, where π_c is the proportion of compliers in the population and $\delta_{\bar{c}}$ is the average treatment effect for noncompliers. Solving for δ_c yields

$$\delta_c = [\delta - (1 - \pi_c)\delta_{\bar{c}}]/\pi_c.$$

Under Assumption 1 (SUTVA) and the following additional assumptions, we can relate this quantity to parameters that can be estimated from the data:

Assumption 2 (exclusion restriction): The treatment assignment R is unrelated to the potential outcomes given the treatment received T—exclusion restriction of treatment assignment given treatment received;

Assumption 3 (monotonicity): There are no defiers—monotonicity of treatment assignment and treatment received;

Assumption 4 (nonzero denominator): The expected difference in the proportion of participants receiving new treatment between those assigned to new treatment and those assigned to control is nonzero; and

Assumption 5 (randomization): Treatment assignment is random and there are no missing data.

Under Assumption 2, the treatment effect for never-takers and for always-takers is identically zero, because these participants always receive the same treatment regardless of assignment. If in addition Assumption 3 holds (there are no defiers), then $\delta_{\bar{c}} = 0$, and hence δ_c equals

$$\delta_c^* = \delta/\pi_c.$$

Under Assumption 5, $\bar{y}_1 - \bar{y}_0$ is an unbiased estimate of δ. Let p_{c+a} be the proportion of participants in the treatment group who take the new treatment and p_a be the proportion of participants in the control group who take the new treatment. Then, under Assumption 5, p_{c+a} is an unbiased estimate of the proportion of compliers or always-takers, and p_a is an unbiased estimate of the proportion of always-takers. Hence p_{c+a} is an unbiased estimate of the proportion of compliers, π_c. Thus an approximately unbiased estimate of δ_c is

$$d_c = (\bar{y}_1 - \bar{y}_0)/(p_{c+a} - p_a), \qquad (1)$$

the estimated intent-to-treat effect divided by difference in the proportion of participants who take the

new treatment in the new treatment and control groups. Assumption 4 assures that the denominator in this expression has a non-zero expectation, that is, that $\pi_c > 0$.

In the JOBS II application there are no always-takers, so $p_a = 0$, and we can denote the estimate of the proportion of compliers in the treatment group as $p_{c+a} = p_c$. Equation 1 then reduces to:

$$d_c = (\bar{y}_1 - \bar{y}_0)/p_c, \tag{2}$$

which is the instrumental variable estimator of the complier-average causal effect derived by Bloom (1984). An equivalent expression is:

$$d_c = (\bar{y}_{c1} - \bar{y}_0) + \frac{1 - p_c}{p_c}(\bar{y}_{n1} - \bar{y}_0), \tag{3}$$

where $\bar{y}_{c1}$ and $\bar{y}_{n1}$ are means for compliers and non-compliers in the treatment group, respectively.

This formulation makes explicit assumptions that were hidden in previous derivations. In particular, Bloom (1984) implied that the randomization Assumption 5 is sufficient for Instrumental Variable Estimator 1 to be valid. Under this assumption d_c is an approximately unbiased estimator of the instrumental variable estimand $\delta_c^* = \delta/\pi_c$. But this assumption alone does not imply that $\delta_c = \delta_c^*$, which is required for Bloom's estimator to be a valid estimator of the complier-average causal effect. Assumptions 1–4 are needed (at least on the average) to assure that $\delta_{\bar{c}} = 0$, so that $\delta_c = \delta_c^*$. In particular the exclusion restriction, Assumption 2, plays a key role and is not a consequence of randomization of treatments. In Section 4 we discuss a second advantage of Imbens and Rubin's (1997b) formulation, the fact that it leads to more efficient estimators of the complier-average causal effect than the instrumental variable estimator. First, however, we apply the ideas of this section to the JOBS II prevention trial.

Application to JOBS II

The JOBS II intervention trial (Vinokur et al., 1995) tested the efficacy of a job training intervention in preventing deterioration in mental health as a result of job loss and in facilitating high quality re-employment. The intervention was 5 half-day job search seminars that helped the participants enhance their job search strategies. The control treatment was a self-guided booklet with advice on how to find a job. Respondents were recruited at the four offices of the Michigan Employment Security Commission in southeastern Michigan. A baseline screening questionnaire (T0) was administered to determine eligibility on the following screening criteria: lost job within 13 weeks, reporting to be looking for a job, not on strike or expecting to be recalled, not planning to retire in next 2 years, no preference among the experimental or control interventions, and absence of a very high depression score. On the basis of the T0 data, 2,464 respondents were selected, invited, and randomized to the field study. All the respondents subject to randomization were mailed a pretest (T1) questionnaire. A total of 1,801 respondents who returned their T1 questionnaire were enrolled in the study, with 552 and 1,249 respondents in the control and experimental groups, respectively. Follow-up questionnaires were mailed to these respondents 6 weeks (T2), 6 months (T3), and 2 years (T4) after the week of the intervention seminar. Demographic variables and measures on depression, financial strain, assertiveness, risk, distress, role and emotional functioning, job search efficacy, self-esteem, mastery, and re-employment were obtained or constructed from the questionnaires.

Noncompliance arises here in the form of a substantial rate (46%) of no-shows for the intervention seminar among the experimental group; receipt of the self-guided booklet was defined to establish compliance in the control group, and hence all participants assigned to that treatment were deemed to have complied with it. If seminars were available to all study participants it would be possible to consider always-takers who attend the seminar regardless of assignment, and defiers who fail to attend the seminars if assigned to them but take the seminars if assigned to the control group. However, because participants randomized to the control group were not allowed to attend the seminars, defiers are forced to act like never-takers because they cannot defy the control assignment, and always-takers are forced to act like compliers because as controls they cannot obtain the intervention. Thus, there is no way to distinguish defiers from never-takers or compliers from always-takers, and analytically nothing is gained from making the distinctions—they are like covariates that cannot be entered into a regression because they are never recorded. Thus we treat always-takers as compliers and defiers as never-takers, and define $C_i = 1$ if the ith participant is a complier and $T_i(R_i) = R_i$ for $R_i = 0, 1$, and $C_i = 0$ if the ith participant is a never-taker and $T_i(R_i) = 0$ for $R_i = 0, 1$.

Note that participants with $R_i = T_i = 1$ were as-

signed and received the intervention and thus are known to be compliers ($C_i = 1$). Participants with $R_i = 1$, $T_i = 0$ were assigned to the intervention and received the control treatment and thus are known to be never-takers ($C_i = 0$). Participants who were assigned and thus received the control treatment ($R_i = T_i = 0$) may or may not have complied with the intervention if assigned to that group, and hence their compliance status (the value of C_i) is unknown. The Instrumental Variable Estimator 2 of the complier-average causal effect reduces to the difference in mean outcomes for participants assigned to the intervention and those assigned to the control divided by the proportion of compliers in the treatment group.

We estimated the complier-average causal effect for the outcome Y = change in depression score between T0 and T3, defined so that positive values reflect an increase in depression over time. To be consistent with previous analyses of the data (Vinokur et al., 1995), we applied the method separately to two groups defined by values of a baseline risk variable. Specifically, we computed a risk score on the basis of the financial strain, assertiveness, and depression scores; participants with a risk score greater or equal to 1.38 were assigned to the high-risk group, and other participants were assigned to the low-risk group. The cutoff point was chosen to obtain approximately 25% high-risk respondents. For the low-risk group the mean depression changes were .016 in the treatment group and .057 in the control group, and the proportion of compliers in the treatment group is .56, so the estimated complier-average causal effect is (.016 − .057)/.56 = −.073. The corresponding calculation in the high-risk group is (−.457 + .383)/.55 = −.132. These estimates are both in the direction of a positive treatment effect but (as discussed in the next section) are not statistically significant.

Let us consider the validity of the assumptions that underlie this estimator in the JOBS II setting. Assumption 3 of no defiers is reasonable in the JOBS II setting because, as noted above, defiance was not an option—participants who were randomized to the control group had no way of attending the intervention. The nonzero denominator assumption (Assumption 4) is also clearly satisfied in the JOBS II setting.

Other assumptions are more questionable. The randomization assumption, 5, is violated because there are missing values, particularly for later outcomes, which are subject to attrition from the sample. For simplicity we ignore that aspect of the data in this discussion, although in a future article we will expand the analysis to incorporate longitudinal outcomes subject to attrition. Assumption 1 (SUTVA) implies that the potential outcome of a respondent does not depend on the treatment status of other respondents. If the presence of a participant in the seminars affects the outcomes of other respondents in the seminar, the potential outcome of participants in the experimental group would depend on whether this influential participant is selected into the experimental group, and SUTVA would be violated. Although SUTVA is a nontrivial assumption for interventions such as JOBS II that involve a group setting, it appears difficult to correct for violations of this assumption, at least without more information on how participants interacted at the seminars. Methods to measure and correct for violations of this assumption are a topic for future research.

Assumption 2 (exclusion restriction) states that the outcome is independent of the treatment assignment given the actual treatment received. That is, the effect of treatment or no treatment for an individual is the same whether the participant is randomized to the experimental group or to the control group. This assumption is the main condition for the randomization indicator to be an instrumental variable. It would be violated if a participant who was randomized to the intervention and did not comply were demoralized by inability to take advantage of the opportunity, whereas the same person randomized to the control group would be less demoralized because the intervention was never offered. The outcome depression score of that individual might well differ under these two scenarios.

Maximum Likelihood Estimation

Analysis Without Covariates

Under Bloom's (1984) conditions and Assumptions 1–5, the instrumental variable estimator is approximately unbiased for the complier-average causal effect. Imbens and Rubin (1997b) proposed an alternative approach to inference for the complier-average causal effect, which builds a model for the distribution of Y and estimates the parameters by maximum likelihood or Bayesian techniques. Suppose there are no defiers, and let π_n, π_a, and π_c denote, respectively, the proportion of never-takers, always-takers, and compliers in the population, $\pi_n + \pi_a + \pi_c = 1$. Assume that the distribution of Y is normal with variance σ^2 and mean μ_n for never-takers, μ_a for always-takers, μ_{c0} for compliers assigned to the control

group, and μ_{c1} for compliers assigned to the treatment group. The likelihood based on the observed data then has the form:

$$L(\theta|data) \alpha \prod_{i\in\{R_i=1,T_i=0\}} \pi_n g(y_i|\mu_n, \sigma^2) \times \prod_{i\in\{R_i=0,T_i=1\}} \pi_a g(y_i|\mu_a, \sigma^2) \times \prod_{i\in\{R_i=1,T_i=1\}} [\pi_c g(y_i|\mu_{c1}, \sigma^2) + \pi_a g(y_i|\mu_a, \sigma^2)] \times \prod_{i\in\{R_i=0,T_i=0\}} [\pi_c g(y_i|\mu_{c0}, \sigma^2) + \pi_n g(y_i|\mu_n, \sigma^2)],$$

where $\theta = (\pi_n, \pi_a, \pi_c, \mu_{c0}, \mu_{c1}, \mu_n, \mu_a, \sigma^2)$ is the set of parameters in the model, and $g(y|\mu, \sigma^2)$ denotes the probability density of a normal distribution with mean μ and variance σ^2. Maximum likelihood estimates of the parameters are obtained by maximizing this function with respect to the parameters θ.

In the JOBS II setting there are no always-takers, so $\pi_a = 0$ and the likelihood simplifies to:

$$L(\theta|data) \alpha \prod_{i\in\{R_i=1,T_i=0\}} \pi_n g(y_i|\mu_n, \sigma^2) \times \prod_{i\in\{R_i=1,T_i=1\}} \pi_c g(y_i|\mu_{c1}, \sigma^2) \times \prod_{i\in\{R_i=0,T_i=0\}} [\pi_c g(y_i|\mu_{c0}, \sigma^2) + \pi_n g(y_i|\mu_n, \sigma^2)], \quad (4)$$

where $\pi_c + \pi_n = 1$. Maximum likelihood estimates can be quite easily computed by means of the EM algorithm (Dempster, Laird, & Rubin, 1977; Little & Rubin, 1987), treating the compliance indicators C_i for participants in the control group as missing data. (See the Appendix for details.) The complier-average causal effect is then estimated by

$$\delta_{\hat{c}} = \hat{\mu}_{c1} - \hat{\mu}_{c0},$$

where $\hat{\mu}_{c0}$ and $\hat{\mu}_{c1}$ are maximum likelihood estimates of μ_{c0} and μ_{c1}. Imbens and Rubin (1997b) noted that Instrumental Variable Estimator 2 can be obtained by subtracting the distribution of outcomes for control noncompliers from the combined distribution of outcomes for control compliers and noncompliers, ignoring the information that the distributions for control compliers and noncompliers must be positive. Maximum likelihood estimation takes into account the fact that these distributions are positive and for large samples yields fully efficient estimates for the parameters under the assumed model. Imbens and Rubin also discussed Bayesian inference for their model using the Gibbs' sampler (Tanner, 1996), which yields improved inferences for small sample sizes.

We now apply this maximum likelihood method to the JOBS II data. Table 1 and 2 show maximum likelihood estimates of θ for the outcome measure Y = change in depression score between T0 and T3 for this model, applied to low-risk and high-risk samples. Standard errors are computed using the bootstrap (see Appendix). For the low-risk group (Table 1), mean changes in depression scores are minor, and the complier-average causal effect (labeled *CACE*) is negative, suggesting a positive effect of treatment, but not significantly different from 0. For the high-risk group (Table 2), the mean changes in depression scores are negative, reflecting a regression to the mean because initial depression was high for these participants. The complier-average causal effect is negative and larger than for the low-risk group, but still not significantly different from 0.

Table 3 is a comparison of the estimated complier-average causal effect from the maximum likelihood and instrumental variable approaches. We computed standard errors using the bootstrap and for the instrumental variable method took into account sampling variability in the proportion of compliers omitted in Bloom's (1984) article. In this application the results of the two methods are very similar. The gain in efficiency of the maximum likelihood method is not necessarily apparent from applications to two specific data sets, but there is a reason for the similarity of the results here. As noted above, the important difference between the two methods is that maximum likelihood,

Table 1
Estimated Means From Normal Model Without Covariates, Low-Risk Group

	Treatment assignment					
	Control ($R_i = 0$)		Treatment ($R_i = 1$)		CACE	
Compliance status	*M*	*SE*	*M*	*SE*	*M*	*SE*
Noncomplier ($C_i = 0$)	.061	.043	.061	.043		
Complier ($C_i = 1$)	.055	.063	−.019	.034	−.074	.071

Note. CACE = complier-average causal effect.

Table 2
Estimated Means From Normal Model Without Covariates, High-Risk Group

	Treatment assignment					
	Control ($R_i = 0$)		Treatment ($R_i = 1$)		CACE	
Compliance status	*M*	*SE*	*M*	*SE*	*M*	*SE*
Noncomplier ($C_i = 0$)	−.391	.063	−.391	.063		
Complier ($C_i = 1$)	−.377	.135	−.513	.050	−.136	.147

Note. CACE = complier-average causal effect.

unlike the instrumental variable method, takes into account the fact that the fitted distributions of the outcome for control compliers and noncompliers must be positive. This constraint is important when (a) the number of noncompliers in the treatment group is high, and (b) the sample mean outcome in the control group differs substantially from the sample mean outcome for noncompliers in the treatment group—if these two means are equal, the instrumental variable and maximum likelihood estimates are equal. In our application the noncompliance rate is quite substantial, but the sample mean changes in depression for the controls and treatment noncompliers are in fact very similar (.06 vs. .06 for the low-risk sample and −.38 vs. −.39 in the high-risk sample). These facts account for the similarity of the estimates in our application.

Analysis With Covariates

We now extend the model of the previous section to allow the mean outcome Y and the probability of compliance π to depend on covariates. We start from the model for Y of the previous section, which we rewrite as

$$y_i = \beta_0 + \beta_C C_i + \beta_{CR} C_i R_i + \epsilon_i,$$

where $\mu_n = \beta_0$, $\mu_{c0} = \beta_0 + \beta_C$, $\mu_{c1} = \beta_0 + \beta_C + \beta_{CR}$, β_{CR} is the complier-average causal effect of treatment, and ϵ_i is a random error normally distributed with mean 0 and variance σ^2. The absence of a main effect for R_i, *say* β_R, in this model is a consequence of exclusion restriction Assumption 2, which implies that the mean outcome given absence of treatment, μ_n, is the same for controls and for noncompliers to treatment.

Let X_Y denote a vector of covariates predictive of the outcome Y. The expanded model for Y is given by:

$$y_i = \beta_0 + \beta_C C_i + \beta_{CR} C_i R_i + \beta_X^T X_{Yi} + (\beta_{CX}^T X_{Yi}) C_i + (\beta_{CRX}^T X_{Yi}) C_i R_i + \epsilon_i, \quad (5)$$

where X_{Yi} is the set of values of X_Y for participant i. In this model, $\beta_0 + \beta_X^T X_Y$ is the mean outcome for noncompliers with covariates X_Y; $\beta_C + \beta_{CX}^T X_Y$ is the difference in mean outcome between control compliers and noncompliers, that is, the effect of compliance; and $\beta_{CR} + \beta_{CRX}^T X_Y$ is the difference in mean outcome between treatment and control compliers, that is, the complier-average causal effect. Exclusion restriction Assumption 2 results in the omission of coefficients involving R but not C in this model. When the coefficients β_X, β_{CX}, and β_{CRX} in this model are all set to 0 we obtain the previous model with no covariates.

The probability of compliance in the model without covariates was a constant π_c. In our extended model the log odds of this probability is expressed as a linear function of the covariates, that is:

$$\log\left(\frac{pr(C_i = 1)}{1 - pr(C_i = 1)}\right) = \gamma_0 + \gamma_X^T X_{Ci}, \quad (6)$$

where X_C is a vector of covariates predictive of compliance and X_{C_i} is the value of X_C for Participant i. Maximum likelihood estimates for the model defined by Equations 5 and 6 can be computed using the EM algorithm, and standard errors can be computed using the bootstrap. More details are provided in the Appendix.

We applied this model to the JOBS II data using covariates measured at baseline. Table 4 lists covariate summary statistics for the control and treatment groups, and for the treatment group, further classified

Table 3
Estimated Complier-Average Causal Effects From Maximum Likelihood and Instrumental Variable Methods

	Maximum likelihood		Instrumental variable	
Risk group	Estimate	*SE*	Estimate	*SE*
Low	−.074	.071	−.073	.077
High	−.136	.147	−.136	.132

Table 4
Means and Standard Deviations of JOBS II Covariates by Treatment and Compliance Status

Covariate	Control group ($R_i = 0$)		Treatment group ($R_i = 1$)					
			All		Noncompliers		Compliers	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
			High-risk group					
Depression at T0	2.49	0.29	2.43	0.30	2.44	0.31	2.42	0.30
Age	36.17	9.75	36.79	10.06	33.31	9.57	39.68	9.56
Risk	1.69	0.19	1.67	0.22	1.68	0.22	1.67	0.21
Motivation to attend	5.32	0.80	5.34	0.82	5.14	0.81	5.50	0.79
School grade completed (2-item)	13.34	1.98	13.38	2.05	12.89	1.90	13.79	2.08
Assertiveness at T0	3.03	0.88	3.09	0.93	3.24	0.92	2.96	0.92
Income	6.80	3.77	6.32	3.83	5.70	3.43	6.84	4.06
Non-married (indicator)	0.58	0.49	0.64	0.48	0.63	0.48	0.65	0.48
Economic hardship at T0	3.47	0.95	3.68	0.82	3.79	0.84	3.60	0.80
Non-White (indicator)	0.18	0.39	0.20	0.40	0.25	0.43	0.15	0.36
			Low-risk group					
Depression at T0	1.47	0.34	1.51	0.35	1.50	0.35	1.51	0.35
Age	36.55	10.72	36.49	10.60	34.57	10.34	37.98	10.58
Risk	0.90	0.27	0.92	0.26	0.90	0.27	0.92	0.25
Motivation to attend	5.11	0.88	5.19	0.89	4.97	0.92	5.36	0.82
School grade completed (2-item)	13.82	1.93	13.79	1.96	13.50	1.92	14.02	1.97
Assertiveness at T0	3.70	0.86	3.69	0.85	3.73	0.85	3.65	0.85
Income	7.52	4.22	7.58	4.42	6.90	4.40	8.10	4.36
Non-married	0.50	0.50	0.52	0.50	0.55	0.50	0.50	0.50
Economic hardship at T0	2.71	0.92	2.68	0.90	2.67	0.89	2.69	0.91
Non-White	0.18	0.39	0.23	0.42	0.25	0.43	0.21	0.41

Note. T0 = baseline.

into noncompliers and compliers. Summary statistics for the control and treatment groups are very similar, reflecting the fact that assignment was randomized and nonresponse bias is minor. Summary statistics for compliers and noncompliers show differences that reflect the nonrandom nature of compliance, with compliers being older, more motivated to attend, more educated, and having higher incomes.

Covariates X_Y found to be significantly related to Y are risk and depression at T0 for high-risk individuals, and marital status (indicator for not married) and depression at T0 for low-risk individuals. The covariates X_C found to be significant in predicting the compliance for high-risk individuals are age, motivation to attend (2-item scale), school grade completed, assertiveness at T0, marital status (indicator for not married), economic hardship at T0, and race (indicator for non-Whites), whereas age, school grade completed, motivation to attend (2-item scale), and income are significant in predicting compliance for low-risk individuals.

The full model of Equation 5 was fitted, and a reduced model with $\beta_{CX} = \beta_{CRX} = 0$ assumes that both the effect of compliance and the complier-average causal effect of the treatment are constant for all values of the covariates. On the basis of a likelihood ratio test, the fit of the full model did not add significantly to the fit of the reduced model, suggesting that the assumptions of the reduced model are consistent with the data. On the other hand, the reduced model was found to fit the data significantly better than the model of the previous section without covariates. Parameter estimates for the reduced models applied to the low-risk and high-risk groups are shown in Table 5, together with bootstrap standard errors.

The parameter estimates of the reduced model indicate that for both groups depression at T0 had a strong negative association with change in depression between T0 and T3, as might be expected. The estimated effects of compliance are not significant in either group. Being married had a significant protective effect in the low-risk group. The complier-average causal effect estimate, adjusted for the covariates, is

Table 5
Estimates From Reduced Models for Y

Parameter	Low-risk group		High-risk group	
	Estimate	*SE*	Estimate	*SE*
Intercept (β_0)	0.951	0.102	1.632	0.293
Compliance (β_C)	−.044	.071	.179	.158
CACE (β_{CR})	−.027	.064	−.309	.136
Not married (β_M)	.114	.036	—	—
Risk score (β_S)	—	—	.911	.256
Depression at T0 (β_D)	−.631	.060	−1.462	0.173
σ^2	.287	.027	.506	.037
ln(likelihood)		−924.336		−729.278
LRT (vs. full model)		1.514		4.312
p value of LRT		.824		.365

Note. CACE = complier-average causal effect; T0 = baseline; LRT = likelihood ratio test. Dashes represent variables omitted from the model.

−0.027 (SE = 0.064, *ns*) in the low-risk group and −0.309 (SE = 0.136, $p < .05$) in the high-risk group. Thus there is evidence of a significant treatment effect in the high-risk group after adjusting for covariates.

Tables 6 and 7 are summaries of the mean changes in depression for high- and low-risk respondents by treatment and compliance status, adjusted for covariates. These are derived as predicted means from the reduced model with other covariates set at their overall means. For the low-risk group changes from baseline are small and not significant. For high-risk respondents, all three groups had a reduction in depression from baseline. The experimental compliers have the greatest reduction (−.547), followed by noncompliers (−.417) and control compliers (−.238). The difference in the first and third of these estimates is the significant complier-average causal effect noted above. The slight (though nonsignificant) evidence that noncompliers do better than control compliers may indicate that noncompliers have personal resources for the job search that makes the intervention seminars less pertinent for them.

Table 8 shows the parameter estimates of the logistic regression of π_c, the probability of compliance, for the reduced models. The model for the low-risk group predicts that compliance will be higher for respondents who are older, more educated, have higher income, and have greater motivation to attend. For high-risk respondents, compliance is predicted to be greater for participants who are White, unmarried, older, more educated, have less economic hardship, are less assertive, and have greater motivation to attend. These effects were generally in the directions expected.

The maximum likelihood analysis can be compared with the instrumental variable approach in the presence of covariates proposed by Bloom (1984). He fitted an additive linear model for Y as a function of covariates X_{Yi} and two indicator variables, the participation indicator $P_i = R_iC_i$, which takes the value 1 for

Table 6
Adjusted Means by Treatment and Compliance Status, Low-Risk Group

Compliance status	Treatment assignment				CACE	
	Control ($R_i = 0$)		Treatment ($R_i = 1$)		$\hat{\beta}_{CR}$	
	M	*SE*	*M*	*SE*	*M*	*SE*
Noncomplier ($C_i = 0$)	0.064	.038	.064	.038		
Complier ($C_i = 1$)	.020	.056	−.007	.031	−.027	.064

Note. CACE = complier-average causal effect.

Table 7
Adjusted Means by Treatment and Compliance Status, High-Risk Group

Compliance status	Treatment assignment				CACE	
	Control ($R_i = 0$)		Treatment ($R_i = 1$)		$\hat{\beta}_{CR}$	
	M	*SE*	*M*	*SE*	*M*	*SE*
Noncomplier ($C_i = 0$)	−.417	.055	−.417	.055		
Complier ($C_i = 1$)	−.238	.121	−.547	.048	−.309	.136

Note. CACE = complier-average causal effect.

Table 8
Estimates From Logistic Regression of Compliance on Covariates for Reduced Model

	Group			
	Low risk		High risk	
Parameter	Estimate	*SE*	Estimate	*SE*
Intercept	−4.747	0.989	−8.738	1.751
Age	.020	.009	.079	.016
School grade completed	.103	.049	.300	.074
Motivation to attend	.489	.104	.667	.165
Income	.047	.024		—
Assertiveness		—	−.376	.128
Not married		—	.541	.299
Economic hardship		—	−.159	.151
Non-White		—	−.499	.321

Note. Dashes represent variables omitted from the model.

participants who were randomized to treatment and complied, and the no-show indicator $NS_i = R_i(1 - C_i)$, which takes the value 1 for participants who were randomized to treatment and failed to comply. The adjusted estimate of the complier-average causal effect for a participant with covariates X_{Y_i} is then given as

$$d_{c,adj} = \hat{b} + \frac{1 - p_c}{p_c}\hat{c}, \tag{7}$$

where $\hat{b}$, $\hat{c}$ are the estimated regression coefficients for P and NS, respectively. Equation 7 reduces to the instrumental variable estimator of Equation 3 when there are no covariates. Bloom's (1984) derivation of Equation 7 ignores variability in the proportions of participants who comply across the covariates. When his additive model is combined with Assumptions 1–4, the complier-average causal effect estimate for a participant with covariates X_{Y_i} is

$$d_c(X_{Y_i}) = \hat{b} + \frac{1 - p_c(X_{Y_i})}{p_c(X_{Y_i})}\hat{c}, \tag{8}$$

where $p_c(X_{Y_i})$ is an estimate of the compliance rate for participants with $X_Y = X_{Y_i}$. The latter can be computed by a logistic regression of the compliance indicator on covariates for participants in the treatment group. We note that Estimator 8 varies across subgroups with different values of X_{Y_i}, whereas our reduced model estimates a constant complier-average causal effect β_{CR} for all values of the covariates. Bloom's overall Estimator 7 is obtained by averaging 8 over the compliers. The estimates and standard errors for the low- and high-risk groups are similar to the maximum likelihood analyses for β_{CR} from our reduced model, as can be seen in Table 9.

Conclusions

We have described methods for inference about the complier-average causal effect, the causal effect of a treatment restricted to the subpopulation of individuals who comply with the treatment assigned. We applied the methods to assess the intervention effect on change in depression between T0 and T3 in the JOBS II trial. We improved precision and interpretability of the analysis by measuring and including covariates that are predictive of the outcome and of compliance. The complier-average causal effect estimated from the model with covariates suggested a positive effect for the intervention for compliers in the high-risk group.

Clearly the value of these analyses depends on whether the complier-average causal effect is a useful quantity to estimate. In the presence of noncompliance, the complier-average causal effect is arguably more informative than the intention-to-treat effect, which fails to take into account whether participants who are randomized to a treatment actually received it or not. The "as-treated" analysis is also potentially misleading. In particular, under the monotonicity assumption (3), the "as-treated" analysis compares the treatment outcome for compliers and always-takers with the control outcome for compliers and never-takers. In the JOBS II setting there are no always-takers, so this comparison is not comparing like with like unless never-takers are comparable with compliers in their outcomes. The complier-average causal effect divides the population according to compliance status in a given experiment, and this division potentially varies in repetitions of the experiment. Nevertheless, the complier-average causal effect seems to

Table 9
Adjusted Complier-Average Causal Effects for Reduced Model Estimated by Maximum Likelihood and by Instrumental Variable Methods

	Maximum likelihood		Instrumental variables	
Risk	Estimate	*SE*	Estimate	*SE*
Low	−.027	.064	−.039	.073
High	−.309	.136	−.272	.125

us worthy of study, and the analysis of outcomes by compliance status seems to us informative.

Any approach to estimating the complier-average causal effect involves assumptions, and a useful feature of Imbens and Rubin's (1997b) formulation is that it makes explicit assumptions that tend to be hidden in earlier formulations. In particular the assumptions of SUTVA, and of exclusion restriction of treatment assigned given treatment received, are not obviously satisfied in prevention trials such as JOBS II. Methods for addressing violations of the SUTVA and exclusion restriction assumptions are not well developed and are a topic for future research. Along these lines, Imbens and Rubin (1997b) proposed a Bayesian sensitivity analysis to assess the impact of violations of the exclusion restriction.

Another attractive feature of Imbens and Rubin's (1997b) approach is the potential for increased efficiency of estimation over the more traditional instrumental variable approach. The gain arises because the maximum likelihood method is a fully efficient method and in particular exploits the fact that the density functions of the distributions of compliers and noncompliers under the control treatment are positive. As noted in the Application to JOBS II section, this positivity constraint is potentially important when (a) noncompliance in the treatment group is high and (b) the mean outcome in the control group differs substantially from the mean outcome for noncompliers in the treatment group. In our application to JOBS II the first of these conditions was met, but the second was not, and gains from the maximum likelihood approach over the simpler instrumental variable approach were not apparent. Imbens and Rubin (1997b) presented a simulation study with low rates of compliance where gains from the maximum likelihood approach are important, using a nonparametric version of the model described here. More simulation studies comparing the two methods would be useful.

A subtle technical distinction between our model for the complier-average causal effect and the instrumental variable approach of Bloom (1984) is that they imply different additivity assumptions for the complier-average causal effect in the presence of covariates, as noted in the discussion following Equation 8. The additivity assumption implied by our reduced model with $\beta_{CX} = \beta_{CRX} = 0$ seems more natural, although whether it is more realistic in practice is an empirical question.

The maximum likelihood inferential methods described here assume large samples. Bayesian implementations of Imbens and Rubin's (1997b) method have the potential to provide better inferences in small samples. In a paper currently in preparation, (Yau & Little, 1998) we describe Bayesian and maximum likelihood methods for repeated-measures data subject to noncompliance and missing data; the results from those analyses applied to JOBS II data are somewhat different from the results of instrumental variable estimation.

References

Angrist, J., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91,* 444–472.

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

Bowden, R. J., & Turkington, D. A. (1984). *Instrumental variables.* Cambridge, England: Cambridge University Press.

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

Gronau, R. (1974). Wage comparisons: A selectivity bias. *Journal of Political Economy, 82,* 1019–1043.

Heckman, J., & Robb, R. (1985). Alternative models for evaluating the impact of interventions. In J. Heckman & B. Singer (Eds.), *Longitudinal analysis of labor market data* (pp. 156–245). New York: Cambridge University Press.

Holland, P. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81,* 945–970.

Imbens, G. W., & Rubin, D. B. (1997a). Bayesian inference for causal effects in randomized experiments with noncompliance. *Annals of Statistics, 25,* 305–327.

Imbens, G. W., & Rubin, D. B. (1997b). Estimating outcome distributions for compliers in instrumental variables models. *Review of Economic Studies, 64,* 555–574.

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

Price, R. H., & Vinokur, A. D. (1995). Supporting career transitions in a time of organizational downsizing: The Michigan JOBS program. In M. London (Ed.), *Employees, careers and job creation: Developing growth-oriented human resource strategies and programs* (pp. 191–209). San Francisco: Jossey-Bass.

Robins, J. M. (1989). The analysis of randomized and nonrandomized AIDS treatment trials using a new approach

to causal inference in longitudinal studies. In L. Sechrest, H. Freeman, & A. Bailey (Eds.), *Health service research methodology: A focus on AIDS* (pp. 113–159). Washington, DC: U.S. Publich Health Service.

Rubin, D. B. (1974). Estimating causal effects of treatment in randomized and non-randomized 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.

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

Tanner, M. A. (1996). *Tools for statistical inference* (3rd ed). New York: Springer-Verlag.

Vinokur, A. D., Price, R. H., & Caplan, R. D. (1991). From field experiments to program implementation: Assessing the potential outcomes of an experimental intervention program for unemployed persons. *American Journal of Community Psychology, 19,* 543–562.

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

Yau, L. H. Y., & Little, R. J. A. (1998). *Inference for the complier-average causal effect from longitudinal data subject to noncompliance and missing data, with application to a job training assessment of the unemployed.* Manuscript in preparation.

Appendix

Computational Details

Maximum Likelihood Estimation for the Complier-Average Causal Effect Without Covariates

The SAS code for implementing the algorithms described in the Maximum Likelihood Estimation section is available on Roderick J. Little's Web site at http://www.sph.umich.edu/~rlittle and can be modified for other applications. We describe here the basic features of the computations.

For the model in the *Analysis Without Covariates* section, with no covariates, it is assumed that: $y_i \sim N(\mu_n, \sigma^2)$ for never-takers, $y_i \sim N(\mu_{c0}, \sigma^2)$ for control compliers, and $y_i \sim N(\mu_{c1}, \sigma^2)$ for experimental compliers. Substituting the relevant normal densities in Equation 2, the log-likelihood of the observed data given $\theta = (\pi_c, \mu_n, \mu_{c0}, \mu_{c1}, \sigma^2)^T$ is

$$
\begin{aligned}
l(\theta|Y,R,T) = {} & N_{11}\log(\pi_c) + N_{10}\log(1-\pi_c) \\
& + \sum_{i\in\{R_i=1,T_i=0\}} \left[-\frac{1}{2} ln(2\pi\sigma^2) - \frac{(y_i-\mu_n)^2}{2\sigma^2} \right] \\
& + \sum_{i\in\{R_i=0,T_i=0\}} \log\left\{ \frac{1-\pi_c}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(y_i-\mu_n)^2}{2\sigma^2} \right) \right. \\
& \left. + \frac{\pi_c}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(y_i-\mu_{c0})^2}{2\sigma^2} \right) \right\} \\
& + \sum_{i\in\{R_i=1,T_i=1\}} \left[-\frac{1}{2} ln(2\pi\sigma^2) - \frac{(y_i-\mu_{c1})^2}{2\sigma^2} \right],
\end{aligned}
$$

where N_{10} is the number of experimental no-shows, N_{11} is the number of experimental shows, and π_c is the probability of complying. Our problem is to maximize this likelihood function with respect to the parameters θ. A convenient approach is to apply the EM algorithm (Dempster et al., 1977), treating the values of C for participants in the control group as missing data. EM is an iterative algorithm consisting of an E (expectation) step, which can be viewed here as imputing the probability of compliance for participants in the control group, and an M (maximization) step, which maximizes an expected complete-data loglikelihood. The algorithm takes the following form:

1. Form initial estimates of the parameters, say $\theta^{(0)}$.
2. At iteration t, given current estimates $\theta^{(t)}$, the E-step consists of computing the estimated probability of compliance for participants i in the control group:

$$
\begin{aligned}
\omega_i^{(t)} &= Pr\{C_i = 1 | R_i = 0, T_i = 0, Y_i, \theta^{(t)}\} \\
&= \frac{\pi_c^{(t)} \exp -\frac{1}{2\sigma^{(t)2}} (y_i - \mu_{c0}^{(t)})^2}{\pi_c^{(t)} \exp -\frac{1}{2\sigma^{(t)2}} (y_i - \mu_{c0}^{(t)})^2 + (1-\pi_c^{(t)}) \exp -\frac{1}{2\sigma^{(t)2}} (y_i - \mu_n^{(t)})^2}.
\end{aligned}
$$

3. The M-step computes new estimates $\theta^{(t+1)}$ of θ as weighted estimates, with participants in the treatment group being classified according their observed compliance status and participants in the control group classified in the compliance group with weight equal to $\omega_i^{(t)}$ and the noncompliance group with weight equal to $1 - \omega_i^{(t)}$. Specifically, the maximum likelihood estimates are:

$$
\hat{\pi}_c^{(t+1)} = \frac{1}{N} \left(N_{11} + \sum_{i\in\{R_i=0,T_i=0\}} \omega_i^{(t)} \right)
$$

$$
\hat{\mu}_{c1}^{(t+1)} = \frac{1}{N_{11}} \sum_{i\in\{R_i=1,T_i=1\}} y_i
$$

$$
\hat{\mu}_{c0}^{(t+1)} = \frac{1}{N_{c0}^{(t+1)}} \sum_{i\in\{R_i=0,T_i=0\}} \omega_i^{(t)} y_i
$$

$$\hat{\mu}_n^{(t+1)} = \frac{1}{N_{10} + N_{00} - N_{c0}^{(t+1)}} \sum_{i \in \{T_i = 0\}} (1 - \omega_i^{(t)}) y_i$$

$$\hat{\sigma}^{2(t+1)} = \frac{1}{N} [N_{11} \sum_{i \in \{R_i = 1, T_i = 1\}} (y_i - \hat{\mu}_{c1}^{(t+1)})^2 + N_{c0}^{(t)} \sum_{i \in \{R_i = 0, T_i = 0\}} \omega_i^{(t)} (y_i - \hat{\mu}_{c0}^{(t+1)})^2 + (N_{10} + N_{00} - N_{c0}^{(t)}) \sum_{i \in (T_i = 0)} (1 - \omega_i^{(t)})(y_i - \hat{\mu}_n^{(t+1)})^2],$$

where $N_{c0}^{(t)} = \Sigma_{i \in \{R_i = 0, T_i = 0\}} \omega_i^{(t)}$ is the estimated number of compliers in the control group.

The EM algorithm then iterates between Steps 2 and 3 until changes in parameter estimates are negligible.

Maximum Likelihood Analysis With Covariates

The EM algorithm in the presence of covariates is similar to the case in which there are no covariates. The E step at Iteration t computes the expectation of C_i given observed values, covariates, and current parameter values $\theta^{(t)}$ for participants in the control group, namely:

$$\omega_i^{(t)} = Pr\{C_i = 1 | R_i = 0, T_i = 0, Y_i, X_{ci}, X_{Y_i}, \theta^{(t)}\}$$

$$= \frac{\pi_c^{(t)} \exp - \frac{1}{2\sigma^{(t)2}} (y_i - \mu_{c0i}^{(t)})^2}{\pi_c^{(t)} \exp - \frac{1}{2\sigma^{(t)2}} (y_i - \mu_{c0i}^{(t)})^2 + (1 - \pi_c^{(t)}) \exp - \frac{1}{2\sigma^{(t)2}} (y_i - \mu_{ni}^{(t)})^2},$$

where $\mu_{c0i}^{(t)}$ is the predicted mean of y_i for compliers in the control group with covariates X_{Y_i} and parameters $\theta^{(t)}$, and $\mu_{ni}^{(t)}$ is the predicted mean of y_i for noncompliers in the control group with covariates X_{Y_i} and parameters $\theta^{(t)}$.

The M step computes the new estimates $\theta^{(t+1)}$ of θ given C, observed values, and covariates. The estimates are given by weighted logistic regression for the compliance model and weighted least squares for the model for Y, with participants in the experimental group being assigned unit weights and participants in the control group being assigned a weight $\omega_i^{(t)}$ for being a complier ($C_i = 1$) and a weight $1 - \omega_i^{(t)}$ for being a never-taker ($C_i = 0$), where values of $\omega_i^{(t)}$ are given above.

Bootstrap Standard Errors

Standard errors are computed with the bootstrap. Specifically, $B = 100$ random samples of size n are drawn with replacement from the original sample, and the above maximum likelihood estimation procedure is applied to each of these samples. The bootstrap estimate of the variance of an estimate is then the sample variance of the estimates over all the bootstrap samples. In particular, if $\hat{\delta}^{(b)}$ is the estimate of δ from the bth bootstrap sample, then the variance of $\hat{\delta}$ is estimated as $\Sigma_b(\hat{\delta}^{(b)} - \bar{\delta})^2/(B - 1)$, where $\bar{\delta} = \Sigma_b \hat{\delta}^{(b)}/B$.

Received June 20, 1997
Revision received December 22, 1997
Accepted December 26, 1997 ■