Handling Missing Data in Randomized Experiments with Noncompliance

Booil Jo*
Department of Psychiatry & Behavioral Sciences
Stanford University
Stanford, CA 94305-5795
booil@stanford.edu

Elizabeth M. Ginexi†
Center for Family Research
George Washington University

Nicholas S. Ialongo
Department of Mental Health
Johns Hopkins University

(Will be published in *Prevention Science*, 2010)

The current version of the article may slightly differ from the final version that will be published in Prevention Science. The final publication will be available at www.springerlink.com. We appreciate helpful feedback from the Prevention Science Methodology Group. *The study of the first author was supported by MH066319 and MH066247 from the National Institute of Mental Health. †The work of the second author was conducted while she was a postdoctoral student at the George Washington University. Elizabeth M. Ginexi is now at the National Institute on Drug Abuse, Bethesda, Maryland.

Handling Missing Data in Randomized Experiments with Noncompliance

ABSTRACT

Treatment noncompliance and missing outcomes at posttreatment assessments are common problems in field experiments in naturalistic settings. Although the two complications often occur simultaneously, statistical methods that address both complications have not been routinely considred in data analysis practice in the prevention research field. This paper shows that identification and estimation of causal treatment effects considering both noncompliance and missing outcomes can be relatively easily conducted under various missing data assumptions. We review a few assumptions on missing data in the presence of noncompliance, including the latent ignorability proposed by Frangakis and Rubin (1999), and show how these assumptions can be used in the parametric CACE estimation framework. As an easy way of sensitivity analysis, we propose the use of alternative missing data assumptions, which will provide a range of causal effect estimates. In this way, we are less likely to settle with a possibly biased causal effect estimate based on a single assumption. We demonstrate how alternative missing data assumptions affect identification of causal effects, focusing on the complier average causal effect (CACE). The data from the Johns Hopkins School Intervention Study (Ialongo et al., 1999) will be used as an example.

Key words: causal inference; complier average causal effect; latent ignorability; missing at random; missing data; noncompliance.

1 Introduction

In field experiments that are frequently employed in prevention research, it is common to have missing data due to dropout or nonresponse (i.e., failing to provide the data) at followup assessments. Along with the extensive development of statistical methods and analysis tools, proper handling of missing data has become a much more manageable task in recent years. Multiple imputation (Schafer, 1997) and likelihood-based estimation methods (Little & Rubin, 2002) are now considered standard analysis options for handling missing data. Another common complication in field experiments is noncompliance of study participants with the assigned programs or treatments. Although not as widely spread as missing data techniques to handle missing outcomes, analytical strategies of handling noncompliance are also increasingly utilized in prevention research (Stuart et al., 2008). Since Angrist, Imbems, and Rubin (1996) demonstrated the possibility of making causal inference in the presence of noncompliance, variations of the instrumental variable approach such as the complier average causal effect (CACE) estimation method have received much attention in various research fields that involve experiments in naturalistic settings. More recently, Frangakis and Rubin (2002) proposed a broader conceptual framework called "principal stratification," which generalizes the idea laid out in Angrist et al. (1996). Principal stratification has rapidly become an essential framework for causal inference taking into account heterogeneity in intermediate outcomes (also known as mediators in social science) including, but not limited to, compliance behavior.

As one may suspect from the apparently similar nature of noncompliance and nonresponse at followup assessments, the two complications often occur simultaneously in field experiments. How noncompliance and missing outcomes can collectively affect

causal effect estimation has also been studied in the principal stratification framework (Dunn et al., 2003; Frangakis & Rubin, 1999; Jo, 2008a; Jo & Vinokur, in press; Mattei & Mealli, 2007; Mealli et al., 2004; O'Malley & Normand, 2004; Peng et al., 2004), although its history is even shorter than that of the CACE estimation method. If treatment compliance is fully observed, noncompliance does not add complexities in handling missing outcome data. However, in reality, compliance information is almost always incomplete. In field experiments, a control condition often does not involve any particular treatments or programs, and therefore we do not observe compliance under the control condition. Some experiments may employ different treatments or programs for different conditions, and in this case, the observed compliance is mostly not comparable across randomized arms. For example, if complying with a standard program is easier than complying with a new program, observed compliance is unlikely to be comparable across the two programs, which leads to a situation where comparable compliance information is partially missing. Given that, co-occurrence of noncompliance and missing outcomes is a concern in causal treatment effect estimation because causal effects can be differently identified depending on what we assume about the unknown relationship between noncompliance and availability of the outcome data.

Although noncompliance and missing outcomes are common problems in prevention research, statistical methods that address both complications have not been routinely considered in its analysis practice. As pointed out in Frangakis and Rubin (1999), if we impose an assumption that deviates from the true relationship between noncompliance and outcome missingness, the causal treatment effect estimates can be biased. However, given that we do not observe the true relationship between the two, investigating possible bias mechanisms and harnessing plausible ranges of causal treatment effects can be an overwhelming task even for researchers with advanced statistical

skills. One practical way of addressing the complication of having both noncompliance and missing outcomes would be to consider a few scientifically plausible assumptions about the relationship between the two. The use of alternative assumptions will at least provide a range of causal effect estimates and therefore will work as a way of sensitivity analysis. In this paper, we will review a few previously discussed assumptions on missing data in the presence of noncompliance and provide a tutorial on using these assumptions in the parametric CACE estimation framework. In particular, we will use the Mplus program (Muthén & Muthén, 1998 - 2009), which has been increasingly used for parametric CACE estimation.

This paper is organized as follows: Section 2 provides a brief description of the Johns Hopkins School Intervention Study, which will be used as an example throughout the paper. Section 3 defines the complier average causal effect in the presence of both noncompliance and missing outcomes. In Section 4, we discuss alternative missing data assumptions and identification of causal effects. Section 5 shows how the alternative missing data assumptions can be used in estimating CACE using the data from the Hopkins Study. Section 6 provides conclusions. Some readers may find the details provided in Sections 3 and 4 somewhat too technical, although these are the key steps for researchers who intend to actively get involved in the missing data modeling process. We intend to provide minimal, but sufficient information to fully understand the alternative procedures of handling missing data in the presence of noncompliance without going back and forth across related technical papers. Readers who want to first see how these methods can be actually implemented may skip the details and focus on the definition of CACE in Section 3 and definitions of missing data assumptions in Section 4, and then go to the real data application in Section 5.

2 Johns Hopkins PIRC School Intervention Study

We will use the school intervention study conducted by the Johns Hopkins University Preventive Intervention Research Center (PIRC) in 1993-1994 (Ialongo et al., 1999) as an example. The PIRC study was designed to improve academic achievement and to reduce early behavioral problems. First-grade children were randomly assigned to the control or to intervention conditions. Two programs were employed in the study: the Classroom-Centered intervention and the Family-School Partnership (FSP) intervention. In this paper, we will compare the control and the FSP intervention groups. In the FSP condition, parents were asked to implement 66 take-home activities related to literacy and mathematics, whereas no special instructions were given to control condition children's parents.

The intervention was provided over the first-grade school year, following a baseline assessment in the early fall. Various outcomes related to academic achievement and behavioral problems were measured at followup assessments. The outcome that will be analyzed in this paper is shy behavior assessed in the spring of the second grade (18 months from the baseline assessment). The shy behavior was measured by the TOCA-R (Teacher Observation of Classroom Adaptation-Revised; Werthamer-Larsson, Kellam, and Wheeler, 1991), which was designed to assess children's adequacy of performance on core tasks in the classroom as rated by the teacher. The shy behavior is a composite variable that consists of TOCA-R items (scale ranges from 1 to 6) such as friendly to classmates, interact with classmates and teachers, play with classmates, and initiate interactions with classmates. Social interaction/engagement is one of the tasks teachers identified in focus groups in the Woodlawn study (Kelleam et al., 1975) as essential to success in the elementary school classroom. In their study, the term shy behavior

was actually used to describe the maladaptive form of the behavior. Children who are socially disengaged would be less likely to participate in class discussion or seek out the help of their teacher when they do not understand something, which may then lead to academic problems. These shy children would also be less likely to develop healthy peer relations since they failed to adequately interact with their peers.

In the PIRC FSP intervention, many parents failed to complete a sufficient amount of assigned activities (66 total activities) and over-reporting of completion was also expected (parents self-reported). Given that, it was expected that the intervention may not show any desirable effects unless parents had reported a quite high level of completion. When the receipt of intervention treatment is defined as completing about two-thirds of activities (45 out of 66 activities), about 47% of children in the FSP intervention condition properly received the intervention treatment. The trial also suffered from subsequent missing outcomes. The overall response (i.e., providing outcome data) rate was 0.84 (0.88 in the intervention, 0.80 in the control) at the second grade followup assessment. In the intervention condition, the average response rate was 0.91 for those who completed 45 or more activities and 0.85 for those who completed less than 45. Since compliance types could not be observed among individuals in the control condition, how noncompliance with the intervention treatments was related to missing data status at the followup under the control condition was also unknown. We may assume that the same relationship between noncompliance and outcome missing data status we observe under the intervention condition would also apply to the control condition, or we may assume something different. The resulting causal treatment effect estimates may be sensitive to what we assume about this unknown missing data mechanism. The key difficulty here is that these assumptions are not directly testable based on observed information.

3 Complier Average Causal Effect (CACE)

Following the convention originated from Angrist et al. (1996), we will define compliers as individuals who would receive the treatment if offered. Noncompliers include three types of individuals. Never-takers are individuals who would not receive the treatment regardless of whether it is offered. Always-takers are individuals who would always receive the treatment regardless of whether it is offered. Defiers are individuals who would do the opposite of what they are assigned to do. Among these three types of noncompliance, we will focus on the never-taker category in this paper. When individuals are not allowed to access the treatment other than the one they are assigned to take, never-taker is the only possible type of noncompliance, which is the case in the JHU PIRC trial. In this trial, students and parents assigned to the control condition did not have an access to the home learning activity materials.

In line with the JHU PIRC trial, we will consider a randomized experiment where individuals are assigned to one of the two conditions. The assignment status $Z_i = 1$ if individual i ($i = 1, ..., N$) is randomly assigned to the intervention, and $Z_i = 0$ if assigned to the control condition. The observed treatment receipt status $S_i = 1$ if individual i received the treatment (in JHU PIRC, this means completing at least 45 intervention activities), and $S_i = 0$ otherwise. The response indicator $R_i = 1$ if i provides outcome information at the second year followup assessment, and $R_i = 0$ if not. Note that Z_i, S_i, R_i are always observed. The shy behavior outcome Y_i is observed if $R_i = 1$, and unobserved if $R_i = 0$.

Let $Y_i(1)$ denote the potential outcome for individual i when assigned to the intervention condition, and $Y_i(0)$ when assigned to the control condition. Then, the effect of treatment assignment can be defined as $Y_i(1) - Y_i(0)$. This definition considers

potential values of posttreatment outcomes under all compared conditions. This way of defining treatment effects is often referred to as the potential outcomes approach (Holland, 1986; Neyman, 1923; Rubin, 1978, 1980). The individual-level causal effect $Y_i(1) - Y_i(0)$ cannot be identified because an individual cannot be assigned to both the treatment and control conditions. However, the causal effect of treatment assignment can be identified at the average level under certain assumptions. The overall average causal effect, also known as the intention to treat (ITT) effect is defined as $\mu_1 - \mu_0$, where μ_1 is the population mean potential outcome under the treatment and μ_0 under the control condition. The key advantage of using the concept of potential outcomes is that underlying assumptions necessary for causal interpretation can be explicitly clarified. This becomes more critical when identifying causal treatment effects that vary across different intermediate posttreatment outcome values such as compliance status because of increased complexity. In recent years, principal stratification (Frangakis & Rubin, 2002) has emerged as a popular method of approaching causal inference with intermediate outcomes in the potential outcomes framework (Jo, 2008b). Principal stratification refers to classification of individuals on the basis of potential values of intermediate outcomes under all treatment conditions that are compared.

Let $S_i(1)$ denote the potential treatment receipt status for individual i when $Z = 1$, and $S_i(0)$ when $Z = 0$. Since individuals were prohibited from receiving a different treatment than the one that they were assigned to, only two compliance types are possible based on Z and S. The principal strata membership, or latent compliance type $C_i = c$ if individual i would receive the treatment when assigned to the intervention condition, and $C_i = n$ if i would not receive the treatment regardless of the intervention

assignment. That is,

$$C_i = \begin{cases} c\,(\text{complier}) & \text{if } S_i(1) = 1, \text{ and } S_i(0) = 0 \\ n\,(\text{never-taker}) & \text{if } S_i(1) = 0, \text{ and } S_i(0) = 0, \end{cases}$$

which implies that C_i is observed if individual i is assigned to the intervention condition, but unobserved if assigned to the control condition.

First, let us define CACE without worrying about the missing data indicator R_i. Let $Y_i(1, S_i(1))$ denote the potential outcome for individual i when assigned to the intervention condition, and $Y_i(0, S_i(0))$ when assigned to the control condition. Then, the effect of treatment assignment can be defined as $Y_i(1, S_i(1)) - Y_i(0, S_i(0))$. In particular, the effect of treatment assignment for complier i is defined as $Y_i(1, S_i(1) = 1) - Y_i(0, S_i(0) = 0)$. This effect cannot be identified because an individual cannot be assigned to both the treatment and control conditions, but can be identified at the average level under certain assumptions. That is, the complier average causal effect (CACE) is defined at the population average level as

$$CACE = \mu_{1c} - \mu_{0c}, \tag{1}$$

where μ_{1c} is the population mean potential outcome for compliers when assigned to the treatment condition and μ_{0c} when assigned to the control condition. Some introductory materials on CACE can be found in several places (e.g., Jo, 2002a; Jo et al., 2008; Little & Yau, 1998).

Recall that, compliance type C_i is observed in the intervention condition, but unobserved in the control condition. In other words, in equation (1), μ_{1c} is observed from the intervention group data, but μ_{0c} is not observed from the control group data. Instead, we observe the overall control condition mean. The distribution of the overall population mean potential outcome under the control condition (μ_0) can

be thought of as a mixture of the complier and never-taker distributions. That is, $\mu_0 = (1 - \pi_c)\,\mu_{0n} + \pi_c\,\mu_{0c}$, where π_c is the proportion of compliers in the population, and μ_{0n} is the population mean potential outcome for never-takers under the control condition. From this mixture, $\mu_{0c} = (\mu_0 - (1 - \pi_c)\,\mu_{0n})/\pi_c$. Then, equation (1) can be rewritten as

$$CACE = \mu_{1c} - \frac{\mu_0 - (1 - \pi_c)\,\mu_{0n}}{\pi_c}. \tag{2}$$

The following assumptions are commonly used in identifying the CACE, and we will assume in this paper that these assumptions hold.

• Assumption 1: Random assignment - Individuals are randomly assigned to the intervention ($Z_i = 1$) or to the control ($Z_i = 0$) condition.

• Assumption 2: Stable unit treatment value (SUTVA) - Potential outcomes (including intermediate outcomes such as compliance) for each person are unrelated to the treatment status of other individuals (Rubin, 1978, 1980, 1990). In the JHU PIRC example, the unit of randomization was a classroom. Therefore, it is likely that the level of interaction among individuals across different treatment conditions remains about the same as that observed when the unit of randomization is an individual (Sobel, 2006). However, interaction within the same classroom can be a problem. In principle, this interaction can be statistically handled in the principal stratification context (Frangakis, Rubin, and Zhou, 2002; Jo et al., 2008), although how well it can be handled with small numbers of clusters (e.g., 9 classrooms in each arm in the JHU PIRC example) is still unclear. In this paper, we assume that this interaction is not considerable.

• Assumption 3: Outcome Exclusion Restriction (OER) - There is no effect of treatment assignment for those who would not change their treatment receipt behavior regardless of the treatment assignment status (i.e., never-takers and always-takers).

This assumption also can be thought of disallowing any direct effect of treatment assignment (Jo, 2008b). Although we assume OER in this paper to focus on missing data assumptions, this assumption may be violated if treatment assignment itself has an effect on the outcome (e.g., psychological effect). In the randomized trial setting we consider in this paper, this assumption means that $\mu_{0n} = \mu_{1n}$, where μ_{1n} is the population mean potential outcome for never-takers under the treatment condition.

• Assumption 4: Monotonicity - There are no defiers. This assumption is automatically satisfied if individuals are not allowed to take other treatments than the one they are assigned to take, as in the JHU PIRC trial.

Under the assumptions 1-4 above, the CACE can be expressed as

$$CACE = \mu_{1c} - \frac{\mu_0 - (1 - \pi_c)\,\mu_{1n}}{\pi_c}, \tag{3}$$

where μ_0, μ_{1c}, μ_{1n}, and π_c are all directly estimable (e.g., using sample statistics, or using maximum likelihood estimation as in this paper) from the observed data. In analyzing data from the JHU PIRC trial using Mplus, we will consistently use maximum likelihood estimation.

Let us now consider availability of the outcome data in defining CACE. Let π_1^R denote the proportion of individuals who provide outcome information in the population under the treatment condition, and π_0^R under the control condition. As with the outcome itself, π_0^R can be thought of as a mixture of the complier and never-taker distributions, which cannot be separately observed. That is, $\pi_0^R = (1-\pi_c)\,\pi_{0n}^R + \pi_c\,\pi_{0c}^R$, where π_{0n}^R is the response rate for never-takers under the control condition, and π_{0c}^R for compliers. With this information, the observed average outcome of the control condition is

$$\mu_0^{obs} = \frac{\pi_{0n}^R}{\pi_0^R}\,(1 - \pi_c)\,\mu_{0n} + \frac{\pi_{0c}^R}{\pi_0^R}\,\pi_c\,\mu_{0c}. \tag{4}$$

From (4), μ_{0c} can be written as

$$\mu_{0c} = \frac{\mu_0^{obs}\,\pi_0^R - \mu_{0n}\,\pi_{0n}^R\,(1-\pi_c)}{\pi_0^R - \pi_{0n}^R\,(1-\pi_c)}. \tag{5}$$

Then, using (5), the equation (2) can be rewritten as

$$CACE = \mu_{1c} - \left\{\frac{\mu_0^{obs}\,\pi_0^R - \mu_{0n}\,\pi_{0n}^R\,(1-\pi_c)}{\pi_0^R - \pi_{0n}^R\,(1-\pi_c)}\right\}. \tag{6}$$

where μ_0^{obs}, π_0^R and π_c are directly estimable from the observed data. Under OER, μ_{0n} can be replaced by μ_{1n}, which is also directly estimable from the observed data. However, further restrictions are necessary to identify π_{0n}^R (recall that compliance status is not observed under the control condition), which is necessary to fully identify CACE. In the following section, we will review a few previously suggested assumptions that can be imposed on π_{0n}^R.

4 Missing Data Assumptions

One way to model the missing data mechanism with incomplete compliance information is to apply the standard missing data assumption that the missing data mechanism is ignorable, or data are missing at random (MAR: Little & Rubin, 2002) conditional on observed information, which includes the observed portion of compliance data. Under MAR, the missing-data mechanism is ignorable for likelihood-based inferences. In other words, the missing data mechanism does not even need to be explicitly modeled when likelihood-based estimation methods are used.

• Missing data assumption 1: Missing At Random (MAR) - The probability of outcome being recorded is not associated with the outcome conditional on treatment

assignment and observed treatment receipt status, and covariates. In the current setting, a sufficient restriction to satisfy this condition is that $\pi_{0c}^R = \pi_{0n}^R$, meaning that compliers and never-takers have the same response rate under the control condition, but they may have different response rates under the treatment condition. Under this assumption, missingness is not attributable to unobserved data including unobserved compliance status under the control condition.

Under MAR, $\pi_{0c}^R = \pi_{0n}^R = \pi_0^R$. With MAR and the common assumptions defined in the previous section, equation (6) can be rewritten as

$$CACE_{MAR} = \mu_{1c} - \left\{ \frac{\mu_0^{obs} - \mu_{1n}\,(1 - \pi_c)}{\pi_c} \right\}, \tag{7}$$

where all involved parameters are directly estimable from the observed data, and therefore CACE is fully identified.

However, if missingness is attributable to unobserved data, the missing-data mechanism is nonignorable. In this case, likelihood-based inferences can be sensitive to whether and how the missing-data mechanism is specified in the statistical model. For example, in the PIRC example, a low level of completion of intervention activities by parents may indicate family instability, meaning that these families are more likely to move from place to place (or children are more likely to be sent to live with a relative or placed in foster care) due to financial stress or other reasons related to drug or alcohol problems, and therefore it is harder to locate these parents and their children at follow-up assessments. In other words, response probability may be higher among potentially well-complying families even in the absence of intervention treatments (i.e., control condition). In this case, missingness is attributable to unobserved compliance status in the control condition. In other words, the missing data mechanism is no longer ignorable, and therefore cannot be modeled within the MAR framework.

Frangakis and Rubin (1999) considered a missing data mechanism called "latent ignorability (LI)," which means that potential outcomes and associated potential outcome missing data indicators are independent within each level of the latent compliance variable (as opposed to observed compliance in MAR). This assumption is weaker than the conventional ignorability (MAR) assumption, and therefore allows for more flexibility in missing data modeling. Since LI is weaker than MAR, if LI is violated, MAR is also violated. Under LI, the missing data mechanism may be attributable not only to the observed compliance data, but also to the unobserved compliance data (i.e., the missing-data mechanism can be nonignorable). However, the LI assumption itself is not sufficient to identify CACE (LI is weaker than MAR). The following two are the examples of missing data assumptions that can be imposed under LI to identify CACE.

• Missing data assumption 2: Response Exclusion Restriction (RER) - For never-takers, the probability of outcome being recorded is not affected by treatment assignment status. In the current setting, this assumption implies that $\pi_{1n}^R = \pi_{0n}^R$ (Frangakis & Rubin, 1999). Some deviation from *RER* is expected in the PIRC trial. Poorly complying families might have felt some benefit from the intervention and might have felt more obliged to provide information at followup assessments than families in the control condition who would have complied poorly if the intervention had been offered. It is also possible that these families might have been demoralized by failing to comply with the intervention and might have provided data less than families in the control condition who would have complied poorly if the intervention had been offered.

With RER and the common assumptions defined in the previous section, equation (6) can be rewritten as

$$CACE_{RER} = \mu_{1c} - \left\{ \frac{\mu_0^{obs}\,\pi_0^R - \mu_{1n}\,\pi_{1n}^R\,(1-\pi_c)}{\pi_0^R - \pi_{1n}^R\,(1-\pi_c)} \right\}, \tag{8}$$

where all involved parameters are directly estimable from the observed data.

• Missing data assumption 3: Stable Complier Response (SCR) - For compliers, the probability of outcome being recorded is unaffected by treatment assignment status. In other words, whether compliant study participants provide the outcome data would not change regardless of intervention assignment. This assumption has been previously considered as a complier version of the response exclusion restriction (Mealli et al., 2004). In the current setting, this assumption implies that $\pi_{0c}^R = \pi_{1c}^R$. In the JHU PIRC trial, compliers' response behavior is likely to be stable regardless of treatment assignment because good compliance is an indicator of family stability. However, this assumption can be violated if compliers provide the outcome data more when assigned to the treatment condition than when assigned to the control condition because they feel benefited from the intervention and feel obliged to provide the data.

Under SCR, $\pi_{0c}^R = \pi_{1c}^R$. With SCR and the common assumptions defined in the previous section, equation (6) can be rewritten as

$$CACE_{SCR} = \mu_{1c} - \left\{ \frac{\mu_0^{obs}\,\pi_0^R - \mu_{1n}\,(\pi_0^R - \pi_{1c}^R \pi_c)}{\pi_{1c}^R \pi_c} \right\}, \tag{9}$$

where all involved parameters are directly estimable from the observed data.

In this paper, we focus on identification of CACE by directly imposing restrictions on the relationship between outcome missingness and noncompliance. However, it is also possible to identify CACE relying on auxiliary information such as from covariates (Jo, 2002b) and proper priors (Hirano et al., 2000), which we do not cover in this paper. For example, in Emsley, Dunn, and White (in press), causal effects are identified under LI imposing restrictions on the effects of covariates on the outcome instead of imposing restrictions on the relationship between outcome missingness and noncompliance (they also provide Mplus input files for the analyses).

5 Application to the PIRC Study

In this section, the three missing data assumptions discussed in the previous section will be applied to the PIRC data to identify CACE. The four common assumptions (random assignment, SUTVA, OER, monotonicity) will be consistently assumed in this process. As shown in the previous section, the models discussed here can be easily estimated using the instrumental variable approach if covariates are not present. The difference between the instrumental variable and maximum likelihood estimates is often minimal, although there are situations where one method performs better than the other (Jo, 2010). In principle, it is possible to incorporate covariates in the instrumental variable framework (Bloom, 1984; Little & Yau, 1998). However, the estimation procedure can be quite cumbersome even without considering missing outcome data, and little is known about how the method works in simultaneously handling noncompliance and missing data when covariates are present. We will use a maximum likelihood estimation approach, which is known to be often more efficient than the instrumental variable approach in estimating the CACE (Imbens & Rubin, 1997; Little & Yau, 1998), in particular when covariates are included in the estimation. To carry out maximum likelihood estimation of CACE, the Mplus program (Muthén & Muthén, 1998-2009) is used. Mplus input setups used for actual PIRC data analyses are provided in Appendix A. Corresponding Mplus output files are provided in Appendix B. For readers who are interested in hands-on experience, we also provide an artificial data set that can be analyzed using the same input files provided in Appendix A. See Appendix C for more details regarding the artificial data.

Let us assume that the following expression represents the true model for the shy behavior outcome at the second grade followup for individual i. Two dummy variables

are used to represent compliers and never-takers. That is, $c_i = 0$ and $n_i = 1$ for a never-taker, and $c_i = 1$ and $n_i = 0$ for a complier. A continuous outcome variable Y for individual i with compliance status c_i and n_i can be expressed as

$$Y_i = \alpha_n^Y \, n_i + \alpha_c^Y \, c_i + \gamma_n^Y \, n_i \, Z_i + \gamma_c^Y \, c_i \, Z_i + \lambda^Y \, x_i + \epsilon_i, \tag{10}$$

where α_n^Y is an outcome intercept for never-takers, and α_c^Y is an outcome intercept for compliers. To represent CACE, we will use γ_c^Y. The effect of treatment assignment on never-takers is γ_n^Y, which is zero under the common assumption OER. A vector of covariates x includes baseline shy behavior, gender, parent's health, and ethnicity. We will assume that covariate effects λ^Y are the same for never-takers and compliers, but this is not an essential assumption and can be relaxed. Alternative ways of imposing restrictions on covariates in identifying CACE can be found, for example, in Jo (2002b). We assume a normally distributed residual ϵ_i with zero mean and variance σ^2. However, in this paper, we do not use normality as an assumption that is central to identification of CACE.

Since we will also model the missing data mechanism, let us also assume the following true model for the missing outcome indicator at the second grade followup for individual i. In the logit scale, the probability of outcome Y being recorded for individual i can be expressed as

$$logit(\pi_i^R) = \alpha_n^R \, n_i + \alpha_c^R \, c_i + \gamma_n^R \, n_i \, Z_i + \gamma_c^R \, c_i \, Z_i + \lambda^R \, x_i, \tag{11}$$

where α_n^R is an logit intercept for never-takers, and α_c^R for compliers. The effect of treatment assignment on outcome missingness is γ_n^R for never-takers, γ_c^R for compliers. We will assume that covariate effects λ^R are the same for never-takers and compliers, but this assumption can be relaxed.

The third component of modeling CACE is the relationship between compliance status and pretreatment covariates. The logistic regression of compliance on covariates is described as

$$\begin{aligned} P(c_i = 1 \mid x_i) &= \pi_{ci}, \\ P(c_i = 0 \mid x_i) &= 1 - \pi_{ci}, \\ logit(\pi_{ci}) &= \beta_0 + \beta_1 x_i, \end{aligned} \tag{12}$$

where π_{ci} denotes the probability of being a complier, β_0 is a logit intercept, and β_1 is a vector of logit coefficients, which represent the association between compliance and pretreatment covariates.

First, we estimate CACE assuming the conventional MAR assumption. This procedure can be carried out in Mplus without explicitly modeling the relationship between outcome missing data status and noncompliance. The Mplus input file for this set-up is shown in Appendix A.1. In this input, the outcome (shy6) is regressed on treatment assignment (Z) and four covariates (shy0, male, health, black). The effect of treatment assignment is freely estimated for the complier class ($C\#1$), and fixed at zero for the never-taker class ($C\#2$) according to OER. The compliance status (C) is regressed on the same covariates in the logistic regression. The missing data indicator (R) is not modeled at all in this set-up. An alternative set-up for CACE estimation under MAR is shown in Appendix A.2. In this set-up, the missing data indicator (R) is explicitly modeled although that is not necessary. The missing data indicator (R) is regressed on treatment assignment (Z) and covariates in the logistic regression. The effect of treatment assignment on the missing data indicator (R ON Z) is allowed to vary across compliance classes, but the logit intercept (R$1) is not, meaning that the response rate is the same across compliance classes under the control condition conditional on covariates. We present this set-up to show how the MAR

assumption can be expressed in the Mplus framework. In this way, how the missing data mechanism is differently modeled across different missing data assumptions can be easily compared. The two Mplus setups shown in Appendix A.1 and Appendix A.2 result in the identical estimates.

Second, we estimate CACE assuming the RER assumption. The Mplus input file for this set-up is shown in Appendix A.3. Compared to the MAR set-up shown is Appendix A.2, a slight change is made in terms of the missing data model. In this input, the effect of treatment assignment on the missing data indicator (R ON Z) is freely estimated for the complier class, but fixed at zero for the never-taker class. In other words, the response rate is allowed to vary across assignment arms for the compliers, but are not allowed for the never-takers. The logit intercept (R$1) is freely estimated in both compliance classes, meaning that the response rate can be different across compliance classes conditional on covariates.

Finally, CACE is estimated under the SCR assumption. The Mplus input file for this set-up is shown in Appendix A.4. Whereas the restriction on the treatment assignment effect on the outcome missingness is imposed for never-takers in the RER model, the same restriction is instead imposed for compliers in the SCR model. In this input, the effect of treatment assignment on the missing data indicator (R ON Z) is freely estimated for the never-taker class, but fixed at zero for the complier class.

Table 1 shows the CACE estimates obtained under different missing data assumptions. In addition to the three missing data models discussed in this paper, we also present the result from the complete case only analysis. In Appendix A.1, if we revive the command "USEOBS = (shy6 NE 999)", which is currently commented out, the analysis will be conducted using complete cases only (i.e., excluding cases with missing

outcome data). Although this strategy is often used in practice, excluding cases with missing outcome data is unnecessary for identification of CACE.

In terms of the logistic regression of compliance on covariates, baseline shy behavior, parent's health, and ethnicity (African-American or not) were significant predictors of compliance in the analyses assuming MAR and SCR. Parents with limited health complied less, African-American parents complied less, parents with kids with higher baseline shy behavior complied less with intervention activities. In the analysis assuming RER, parent's health and ethnicity were significant predictors of compliance. In the analysis with complete cases only, only ethnicity was a significant predictor of compliance.

[Table 1]

In the results reported in Table 1, negative values mean positive effects of treatment assignment (i.e., assignment to the treatment condition lowers the level of shy behavior). Given that, let us evaluate the magnitude of CACE estimates in terms of their absolute values. The CACE estimator assuming SCR resulted in the smallest CACE estimate and the estimator assuming RER resulted in the largest CACE estimate. The resulting range of CACE estimates is quite narrow, suggesting that there was a significant positive effect of treatment assignment on the shy behavior outcome at the followup. However, both of these assumptions may be violated and we cannot exclude the possibility that the true CACE is still not captured within the range of CACE estimates we obtained based on considered missing data assumptions. The ideal situation in this case would be that SCR actually yields the lower bound and RER the upper bound for CACE. In particular, the lower bound is often more of concern in prevention studies. As discussed earlier, in the JHU PIRC trial, compliers' response

behavior is likely to be stable, but if not, it is very likely that compliers provide the data more under treatment condition than under the control condition. We can analytically derive how this directional knowledge would affect the CACE estimate, but a quick and easy way would be to empirically estimate CACE with a higher response rate of compliers under the treatment condition. In Appendix A.4, this can be done by changing R ON Z@0 to, for example, R ON Z@1. Whereas the response rate remains the same across treatment arms when this coefficient is fixed at zero (i.e., SCR), the response rate is higher for the treatment condition if we fix this at a positive value in the logistic regression. The absolute value of the CACE estimate with this modified setting is 0.555 (compared to 0.477 under SCR) meaning that the CACE estimate gets larger if the response rate for compliers is higher under the treatment condition. In other words, it is very unlikely that the magnitude of true CACE is smaller than the one estimated under SCR.

6 Conclusion

It was demonstrated in this paper that identification and estimation of causal treatment effects considering both noncompliance and missing outcomes can be relatively easily conducted under various missing data assumptions. In particular, parametric maximum likelihood estimation of CACE can be carried out using the mixture analysis feature embedded in widely used latent variable modeling software such as Mplus (Muthén & Muthén, 1998 - 2009). Given that, there is little reason to settle with a single solution or to exclude cases with missing outcome data when estimating CACE in practice. Nonetheless, the methods we discussed in this paper are more complicated than the standard ways of handling missing data. When we handle outcome missing

data by listwise deletion or in the likelihood-based estimation framework assuming MAR (Little & Rubin, 2002), there is no need to explicitly model the missing data mechanism. In this paper, we discussed missing data assumptions that require explicit modeling. This necessitates more active involvement of researchers in the missing data modeling and identification of causal effects, which has not beed advocated much in the prevention research field.

As discussed in the context of the PIRC trial, we usually do not have one missing data assumption we are absolutely confident about. In fact, we rarely have a clear idea about which assumption is more likely than others. Even if we know the relative plausibility of these assumptions, it is still difficult to predict relative performance of CACE estimators that operate under different assumptions because the bias mechanism can be quite complex when missing data are accompanied by treatment noncompliance (Jo, 2008a). Given the complexities at hand, as a practical way of sensitivity analysis, we proposed the use of alternative missing data assumptions, which will at least provide a range of causal effect estimates. This is already a significant improvement compared to the usual practice of CACE estimation with no sensitivity analysis. However, researchers may go one step further and check if the true CACE is likely to fall within the range of CACE estimates obtained with alternative assumptions. This is possible when researchers are confident about the directionality of missing data assumptions. We showed a simple way of checking bounds in this paper. More elaborate ways of establishing bounds for CACE utilizing all considered assumptions can be found in Jo and Vinokur (in press).

In this paper, we assumed that the exclusion restriction on the outcome (OER) always holds and focused on how different missing data assumptions affect CACE estimation. However, in field experiments where blinding is rarely possible, this as-

sumption may also be violated, requiring sensitivity analysis considering both missing data and outcome assumptions. In the PIRC trial, never-takers can partially receive the treatment if assigned to the treatment condition (recall that never-takers were defined as those who completed less than 45 out of 66 intervention activities), and therefore, plausibility OER is questionable. To maintain simplicity in introducing the use of missing data assumptions, which is already somewhat complicated in the CACE estimation context, we chose to use the consistent outcome assumption (i.e., OER). However, it is possible to conduct sensitivity analysis considering variations of both missing data and outcome assumptions (Jo & Vinokur, in press). This paper also limited its discussion to CACE estimation. However, the methods discussed here can be applied to more general causal inference problems involving intermediate outcomes (not limited to compliance) and outcome missing data.

Appendix A. PIRC Example Mplus Input Files

A.1. CACE Estimation Under MAR

```
TITLE: CACE estimation under MAR
DATA: FILE = ps09jhu.dat;
VARIABLE:NAMES = Z S R shy6 shy0 male health black;
USEV = Z S shy6 shy0 male health black;
! USEOBS = (shy6 NE 999);
CATEGORICAL = S; ! binary compliance indicator (0/1, missing=999);
CLASSES = C(2); !two compliance strata
MISSING = all (999); ! missing values coded as 999;
ANALYSIS:TYPE = MIXTURE;

MODEL:
%OVERALL%
shy6 ON Z shy0-black; !shy behavior at 6 months regressed on
                      !randomization status Z and covariates
C#1 ON shy0-black; !compliance class C regressed on covariates

%C#1%
    [S$1@-15]; !compliers
shy6 ON Z; !compliers' outcome varies across Z

%C#2%
    [S$1@15]; !never-takers
shy6 ON Z@0; !never-takers' outcome is stable across Z (OER)
```

A.2. CACE Estimation Under MAR II

TITLE: CACE estimation under MAR II
DATA: FILE = ps09jhu.dat;
VARIABLE:NAMES = Z S R shy6 shy0 male health black;
USEV = Z S **R** shy6 shy0 male health black;
CATEGORICAL = S; ! binary compliance indicator (0/1, missing=999);
CATEGORICAL = R; ! binary outcome missing indicator for shy6;
CLASSES = C(2); !two compliance strata
MISSING = all (999); ! missing values coded as 999;
ANALYSIS:TYPE = MIXTURE;

MODEL:
%OVERALL%
shy6 ON Z shy0-black; !shy behavior at 6 months regressed on
!randomization status Z and covariates
C#1 ON shy0-black; !compliance class C regressed on covariates
R ON Z shy0-black; !missing data status is related to observed information

%C#1%
[S$1@-15]; !compliers
[R$1] (1); !missing data status stable across C under the control (MAR)
shy6 ON Z; !compliers' outcome varies across Z
R ON Z; !compliers' missing data status varies across Z

%C#2%
[S$1@15]; !never-takers
[R$1] (1); !missing data status stable across C under the control (MAR)
shy6 ON Z@0; !never-takers' outcome is stable across Z (OER)
R ON Z; !never-takers' missing data status varies across Z

A.3. CACE Estimation Under RER

```
TITLE: CACE estimation under RER
DATA: FILE = ps09jhu.dat;
VARIABLE:NAMES = Z S R shy6 shy0 male health black;
USEV = Z S R shy6 shy0 male health black;
CATEGORICAL = S; ! binary compliance indicator (0/1, missing=999);
CATEGORICAL = R; ! binary outcome missing indicator for shy6;
CLASSES = C(2); !two compliance strata
MISSING = all (999); ! missing values coded as 999;
ANALYSIS:TYPE = MIXTURE;

MODEL:
%OVERALL%
shy6 ON Z shy0-black; !shy behavior at 6 months regressed on
                      !randomization status Z and covariates
C#1 ON shy0-black; !compliance class C regressed on covariates
R ON Z shy0-black; !missing data status R is related to observed information;

%C#1%
    [S$1@-15]; !compliers
    [R$1]; !missing data status varies across C under the control
shy6 ON Z; !compliers' outcome varies across Z
R ON Z; !compliers' missing data status varies across Z

%C#2%
    [S$1@15]; !never-takers
    [R$1]; !missing data status varies across C under the control
shy6 ON Z@0; !never-takers' outcome is stable across Z (OER)
R ON Z@0; !never-takers' missing data status is stable across Z (RER)
```

A.4. CACE Estimation Under SCR

```
TITLE: CACE estimation under SCR
DATA: FILE = ps09jhu.dat;
VARIABLE:NAMES = Z S R shy6 shy0 male health black;
USEV = Z S R shy6 shy0 male health black;
CATEGORICAL = S; ! binary compliance indicator (0/1, missing=999);
CATEGORICAL = R; ! binary outcome missing indicator for shy6;
CLASSES = C(2); !two compliance strata
MISSING = all (999); ! missing values coded as 999;
ANALYSIS:TYPE = MIXTURE;

MODEL:
%OVERALL%
shy6 ON Z shy0-black; !shy behavior at 6 months regressed on
                      !randomization status Z and covariates
C#1 ON shy0-black; !compliance class C regressed on covariates
R ON Z shy0-black; !missing data status R is related to observed information;

%C#1%
    [S$1@-15]; !compliers
    [R$1]; !missing data status varies across C under the control
shy6 ON Z; !compliers' outcome varies across Z
R ON Z@0; !compliers' missing data status is stable across Z (SCR)

%C#2%
    [S$1@15]; !never-takers
    [R$1]; !missing data status varies across C under the control
shy6 ON Z@0; !never-takers' outcome is stable across Z (OER)
R ON Z; !never-takers' missing data status varies across Z
```

Appendix B. PIRC Example Mplus Output Files (key model parameter estimates only)

B.1. CACE Estimation Under MAR

	Estimate	S.E.	Est./S.E.	P-Value
Latent Class 1 (complier)				
SHY6 ON				
Z (**CACE**)	**-0.553**	0.229	-2.417	0.016
SHY0	0.228	0.052	4.372	0.000
MALE	0.223	0.112	1.982	0.047
HEALTH	0.305	0.211	1.443	0.149
BLACK	0.019	0.162	0.115	0.908
Intercepts				
SHY6	2.256	0.253	8.908	0.000
Residual Variances				
SHY6	0.924	0.083	11.173	0.000
Latent Class 2 (never-taker)				
SHY6 ON				
Z (OER)	0.000	0.000	999.000	999.000
Intercepts				
SHY6	1.431	0.212	6.764	0.000
Logistic Regression of C on X				
C#1 ON				
SHY0	-0.300	0.151	-1.987	0.047
MALE	0.135	0.290	0.465	0.642
HEALTH	-1.094	0.538	-2.035	0.042
BLACK	-1.044	0.412	-2.536	0.011
Intercepts				
C#1	1.437	0.499	2.878	0.004

B.2. CACE Estimation Under MAR II

	Estimate	S.E.	Est./S.E.	P-Value
Latent Class 1 (complier)				
SHY6 ON				
Z (**CACE**)	**-0.553**	0.229	-2.417	0.016
SHY0	0.228	0.052	4.372	0.000
MALE	0.223	0.112	1.982	0.047
HEALTH	0.305	0.211	1.443	0.149
BLACK	0.019	0.162	0.115	0.908
Intercepts				
SHY6	2.256	0.253	8.908	0.000
Residual Variances				
SHY6	0.924	0.083	11.173	0.000
R (missing indicator) ON				
Z	0.952	0.395	2.409	0.016
SHY0	0.014	0.123	0.114	0.909
MALE	-0.226	0.278	-0.814	0.415
HEALTH	-0.212	0.402	-0.528	0.597
BLACK	0.781	0.324	2.412	0.016
Thresholds (MAR: [R$1] (1) in Input)				
R$1	**-0.902**	0.453	-1.989	0.047
Latent Class 2 (never-taker)				
SHY6 ON				
Z (OER)	0.000	0.000	999.000	999.000
Intercepts				
SHY6	1.431	0.212	6.764	0.000
R (missing indicator) ON				
Z	0.298	0.324	0.920	0.357
Thresholds (MAR: [R$1] (1) in Input)				
R$1	**-0.902**	0.453	-1.989	0.047

B.3. CACE Estimation Under RER

	Estimate	S.E.	Est./S.E.	P-Value
Latent Class 1 (complier)				
SHY6 ON				
Z (**CACE**)	**-0.586**	0.251	-2.332	0.020
SHY0	0.227	0.052	4.368	0.000
MALE	0.225	0.112	2.004	0.045
HEALTH	0.306	0.212	1.441	0.149
BLACK	0.016	0.162	0.098	0.922
Intercepts				
SHY6	2.291	0.277	8.266	0.000
Residual Variances				
SHY6	0.920	0.083	11.033	0.000
R (missing indicator) ON				
Z	1.171	0.537	2.178	0.029
SHY0	-0.003	0.124	-0.025	0.980
MALE	-0.219	0.280	-0.780	0.435
HEALTH	-0.294	0.417	-0.704	0.481
BLACK	0.717	0.343	2.092	0.036
Thresholds				
R$1	-0.757	0.525	-1.442	0.149
Latent Class 2 (never-taker)				
SHY6 ON				
Z (OER)	0.000	0.000	999.000	999.000
Intercepts SHY6	1.437	0.210	6.848	0.000
R (missing indicator) ON				
Z (RER)	0.000	0.000	999.000	999.000
Thresholds				
R$1	-1.276	0.548	-2.328	0.020

B.4. CACE Estimation Under SCR

	Estimate	S.E.	Est./S.E.	P-Value
Latent Class 1 (complier)				
SHY6 ON				
Z (**CACE**)	**-0.477**	0.205	-2.325	0.020
SHY0	0.228	0.053	4.332	0.000
MALE	0.215	0.113	1.899	0.058
HEALTH	0.301	0.210	1.436	0.151
BLACK	0.025	0.161	0.155	0.877
Intercepts				
SHY6	2.179	0.231	9.423	0.000
Residual Variances				
SHY6	0.934	0.082	11.380	0.000
R (missing indicator) ON				
Z (SCR)	0.000	0.000	999.000	999.000
SHY0	0.065	0.128	0.507	0.612
MALE	-0.300	0.300	-0.999	0.318
HEALTH	0.001	0.443	0.002	0.999
BLACK	1.028	0.385	2.670	0.008
Thresholds				
R$1	-1.620	0.516	-3.139	0.002
Latent Class 2 (never-taker)				
SHY6 ON				
Z (OER)	0.000	0.000	999.000	999.000
Intercepts				
SHY6	1.419	0.217	6.533	0.000
R (missing indicator) ON				
Z	0.866	0.413	2.097	0.036
Thresholds				
R$1	-0.021	0.605	-0.035	0.972

Appendix C. Artificial Data Analyses

For readers who are interested in hands-on experience, we provide an artificial data set, which can be obtained from the Prevention Science website (URL to be provided by Prevention Science). The same Mplus input files provided in Appendix A can be used after changing the data file name (i.e., DATA: FILE = **artif.dat**;). The results of the artificial data analyses are provided below in Table 2.

Table 2. Artificial Data: CACE Estimates Under Different Missing Data Assumptions (standard error in parentheses).

CC	MAR	RER	SCR
–0.584	–0.617	–0.632	–0.546
(0.169)	(0.174)	(0.178)	(0.180)

REFERENCES

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

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

Dunn, G., Maracy, M., Dowrick, C., Ayuso-Mateos, J. L., Dalgard, O. S., Page, H., Lehtinen, V., Casey, P., Wilkinson, C., Vasquez-Barquero, J. L., & Wilkinson, G. (2003). Estimating psychological treatment effects from a randomized controlled trial with both non-compliance and loss to follow-up. *British Journal of Psychiatry, 183*, 323-331.

Emsley, R., Dunn, G., & White, I. R. (in press). Mediation and moderation of treatment effects in randomised controlled trials of complex interventions. *Statistical Methods in Medical Research*.

Frangakis, C. E. & Rubin, D. B. (1999). Addressing complications of intention-to-treat analysis in the presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika, 86*, 365-379.

Frangakis, C. E. & Rubin, D. B. (2002) Principal stratification in causal inference. *Biometrics, 58*, 21-29.

Frangakis, C. E., Rubin, D. B., & Zhou, X. H. (2002). Clustered encouragement design with individual noncompliance: Bayesian inference and application to advance directive forms. *Biostatistics, 3, 147-164*.

Hirano, K., Imbens, G. W., Rubin, D. B., & Zhou, X. H. (2000). Assessing the effect of an influenza vaccine in an encouragement design. *Biostatistics, 1*, 69-88.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*, 945-960.

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. & Rubin, D. B. (1997). Bayesian inference for causal effects in randomized experiments with non-compliance. *Annals of Statistics, 25*, 305-327.

Jo, B. (2002a). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods, 7*, 178-193.

Jo, B. (2002b). Estimating intervention effects with noncompliance: Alternative model specifications. *Journal of Educational and Behavioral Statistics, 27*, 385-420.

Jo, B. (2008a). Bias Mechanisms in intention-to-treat analysis with data subject to treatment noncompliance and missing outcomes. *Journal of Educational and Behavioral Statistics, 33*, 158-185.

Jo, B. (2008b). Causal inference in randomized experiments with mediational processes. *Psychological Methods, 13*, 314-336.

Jo, B. (2010). Estimation of principal causal effects using maximum-likelihood and instrumental variable approaches. Manuscript in preparation.

Jo, B., Asparouhov, T., Muthén, B. O., Ialongo, N. S., & Brown, C. H. (2008). Cluster randomized trials with treatment noncompliance. *Psychological Methods, 13*, 1-18.

Jo, B., & Vinokur, A. (in press). Sensitivity analysis and bounding of causal effects with alternative identifying assumptions. *Journal of Educational and Behavioral Statistics*.

Kellam, S. G., Branch, J. D., Agrawal, K. C., & Ensminger, M. E. (1975). *Mental health and going to school: The Woodlawn program of assessment, early intervention, and evaluation*. Chicago: University of Chicago Press.

Little, R. J. A. & Rubin, D. B. (2002). *Statistical analysis with missing data*. New York: John Wiley & Sons

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.

Mattei, A., & Mealli, F. (2007). Application of the principal stratification approach to the Faenza randomized experiment on breast self-examination. *Biometrics, 63*, 437-446.

Mealli, F., Imbens, G. W., Ferro, S., & Biggeri A. (2004). Analyzing a randomized trial on breast self-examination with noncompliance and missing outcomes. *Bio-*

statistics, 5, 207-222.

Muthén, L. K. and Muthén, B. O. (1998-2009). *Mplus user's guide*. Los Angeles: Muthén & Muthén.

Neyman, J. (1923). On the application of probability theory to agricultural experiments. Section 9 translated in *Statistical Science, 5*, 465-480 (1990).

O'Malley, A. J., & Normand, S. L. T. (2004). Likelihood methods for treatment noncompliance and subsequent nonresponse in randomized trials. *Biometrics, 61*, 325-334.

Peng, Y., Little, R. J., & Raghunathan, T. E. (2004). An extended general location model for causal inferences from data subject to noncompliance and missing values. *Biometrics, 60*, 598-607.

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.

Schafer, J. L. (1997). *Analysis of incomplete multivariate data*. CRC press.

Sobel, M. E. (2006). What do randomized studies of housing mobility demonstrate: Causal inference in the face of interference. *Journal of the American Statistical Association, 101*, 1398-1407.

Stuart, E. A., Perry, D. F., Le, H-N, & Ialongo, N. S. (2008). Estimating intervention effects of prevention programs: Accounting for noncompliance. *Prevention Science, 9*, 288-298.

Werthamer-Larsson, L., Kellam, S. G., and 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.

Table 1. JHU PIRC: CACE Estimates Under Different Missing Data Assumptions (standard error in parentheses). CC: complete cases only, MAR: missing at random, RER: response exclusion restriction, SCR: stable complier response.

CC	MAR	RER	SCR
–0.545	–0.553	–0.586	–0.477
(0.229)	(0.229)	(0.251)	(0.205)