Journal of Educational and Behavioral Statistics
Winter 2002, Vol. 27, No. 4, pp. 417–420

Rejoinder by Booil Jo

I sincerely thank Fabrizia Mealli and Donald B. Rubin for their thorough and constructive commentary on the article. I would like to take this opportunity to clarify a few issues on the basis of their comments, which I find very helpful and reasonable. Most of all, I thank discussants for several suggestions related to real data examples. I hope that the readers of this article will note that the analyses in real data examples were focused on illustrating identification of proposed CACE models without involving other concurrent problems such as missing data in outcomes and covariates. Specific points in the four issues raised by discussants are addressed as follow.

Sensitivity of Significant/Not Significant Results

The first issue raised by the discussants is that sensitivity analysis of alternative models seems to be solely based on "significant" and "not significant" results. I agree with the discussants that more information than significance of the results can be provided to better describe sensitivity of the CACE estimate to violation (or relaxation) of the exclusion restriction assumption. I believe that employing a Bayesian paradigm will be very useful in evaluating precision and sensitivity among alternative models, even when CACE models with unique MLE are of concern.

I would like to point out that insensitivity of the CACE estimate in JOBS II is at least partly due to the presence of covariates that are good predictors of compliance (Jo, 2002a). It was not emphasized in this article that covariate information may play an important role not only in models without the exclusion restriction (in identification), but also in models with the exclusion restriction (in reduction of bias). Table 11 shows CACE estimates assuming additive effect of treatment assignment (see Tables 2 and 3 for the full results with nine covariates). It is shown that the CACE estimate is less sensitive (at least in terms of size) to the exclusion restriction in the presence of covariates associated with compliance (e.g., Mot, Age, Assert). Insensitivity of the CACE estimate in JOBS II is also due to the fact that the exclusion restriction is not severely biased. It would be interesting to see how sensitivity of the CACE estimate differs in various examples in different fields.

Dealing with Missing Data by Listwise Deletion

The discussants also suggested a better handling of missing data in covariates and outcomes. As they pointed out, noncompliance and missing data can be indeed considered simultaneously. Instead of assuming missing completely at random (MCAR) (Little & Rubin, 2002), noncompliance and missing data can be simultaneously considered by assuming missing at random (MAR) (Little & Rubin, 2002) as demonstrated in Yau and Little (2001), or by assuming MAR

TABLE 11
JOBS II: Sensitivity of the CACE Estimate to the Exclusion Restriction and Covariate Information

	Exclusion Restriction	
Present covariates	Yes	No
Dep0	–0.297 (0.201)	–0.440 (0.150)
Dep0, Mot, Age, Assert	–0.340 (0.199)	–0.444 (0.140)
All nine covariates	–0.361 (0.191)	–0.451 (0.131)

Note. SE in parentheses.

conditioning on compliance type (latent ignorability), which is less restrictive than the regular MAR assumption (Baker, 2000; Barnard et al., 2002; Frangakis & Rubin, 1999).

The potential relationship between noncompliance and missing data may introduce more complex issues in identifying CACE models, which I did not want to include within the scope of the current article. In fact, in JOBS II, a substantial number of individuals did not respond at follow-up surveys, and observed nonresponse rates in the treatment condition were quite different for compliers and noncompliers, implying potential correlation between noncompliance and nonresponse. To maintain identifiability of CACE models allowing for this correlation (latent ignorability) (Frangakis & Rubin, 1999), the exclusion restriction is imposed not only on outcomes, but also on missing indicators of outcomes (compound exclusion restriction). This situation raises a question of how the exclusion restriction assumption should be tested with an increased number of parameters and increased complexity in bias mechanisms (Jo, in press; Muthén, Jo, & Brown, in press). An interesting next step will be to investigate how proposed alternative assumptions can be applied to build identifiability in this situation, whether resulting model properties lead to reasonable and interpretable estimates, and whether model estimates maintain a practical level of accuracy and precision.

Language that Tends to Confuse Subpopulation Differences and Treatment Effects

This confusion resulted from mechanical interpretation of interaction effects without considering the difference between pretreatment covariates and treatment assignment. In regard to Table 4, the statement "However, high motivation and being married had a negative impact on individuals who would have complied with the intervention if offered, but were assigned to the control condition" (p. 401), needs to be disregarded. In regard to Table 5, the statement "However, being single and less assertive had a negative impact on individuals who would have complied with the intervention if offered, but were assigned to the control condition" (p. 402), needs to be disregarded.

Scientific Rationale for Various Specifications

Finally, the discussants raised the issue of scientific plausibility of the alternative models and the resulting estimates. In some situations, it may be relatively clear which assumptions (therefore which models) are more plausible, possibly based on scientific evidence, previous studies, and experts' opinion. In many other situations, however, plausibility of model assumptions is often questionable. The second best thing one can expect under this uncertainty is to have assumptions that will yield clear and consistent bias mechanisms when violated so that inferences can be made considering possible ranges of bias in parameter estimates. For example, the bias mechanism in CACE models assuming the exclusion restriction is quite straightforward in the absence of covariates. However, in the presence of covariates, the bias mechanism becomes very complicated, which makes interpretation of analysis results very difficult when plausibility of the exclusion restriction is questionable. Similarly, CACE models assuming constant effects of covariates (Model C) involve complex bias mechanisms when the assumption is violated. Therefore, inferences made in these two models depend heavily on the scientific plausibility of model assumptions (i.e., exclusion restriction or constant effects of covariates).

The advantage of CACE models assuming the additive effect of treatment assignment (Model B) is that the bias mechanism is relatively simple when the assumption is violated. The simplicity of the bias mechanism provides a couple of convenient properties in Model B. First, the bias mechanism of compliers is separated from that of never-takers. As a result, regardless of violation of additivity, at least combined (main and interaction) effects of treatment assignment can be estimated without introducing bias due to violation of the exclusion restriction. Second, if there are multiple covariates, the additivity assumption is not completely unverifiable. As shown in Tables 4 and 5, some interaction effects are estimable. On the basis of these two properties, Model B can be creatively used to explore more plausible and parsimonious sets of model assumptions. For example, in the presence of multiple covariates, one can estimate interaction effects for most covariates in each class, and then decide whether additivity or a constant effect assumption is appropriate.

Model B including only compliers and never-takers is useful in more controlled experiments such as JOBS II and Hopkins PIRC. However, in other larger scale experiments or randomized encouragement studies, for practical and ethical reasons, it may be difficult to prohibit study participants assigned to the control condition from receiving treatment (e.g., Barnard et al., 2002; Hirano et al., 2000). As the discussants pointed out, Model B cannot be identified with both never-takers and always-takers when there is only one covariate. One way to build identifiability in this situation is to relax the exclusion restriction for either never-takers or always-takers, but not for both. In doing so, however, Model B loses its unique properties unless there is strong scientific plausibility that the exclusion restriction should hold either for never-takers or for always-takers. I agree with the discussants that it is more difficult to retain interpretability when the exclusion restriction has to be partially imposed in conjunction with the additivity assumption. Here, the information from

multiple covariates comes into play. If there are multiple covariates, the number of subpopulations for which the exclusion restriction can be relaxed increases in Model B. For example, assume there are two binary covariates X_1 and X_2. Compared to the number of parameters and directly estimable population means based on equation (11a), the number of parameters increases by three (i.e., λ_{nX_2}, λ_{aX_2}, λ_{cX_2}), whereas the number of directly estimable population means increases by four (i.e., $\mu_{1n,X_1=0,X_2=1}$, $\mu_{0a,X_1=0,X_2=1}$, $\mu_{00,X_1=0,X_2=1}$, $\mu_{11,X_1=0,X_2=1}$), which results in a just identified model. In principle, Model B with always-takers can be identified with one less degrees of freedom compared to Model B without always-takers. However, whether resulting parameter estimates will retain a practical level of precision is another interesting issue to be studied.

References

Jo, B. (2003). *Sensitivity of causal effects under ignorable and latent ignorable missing-data mechanisms*. Manuscript submitted for publication.

Muthén, B. O., Jo, B., & Brown, C. H. (in press). Discussion on "A principal stratification to broken randomized experiments" by Barnard, Frangakis, Hill, and Rubin. *Journal of American Statistical Association.*

Yau, L. H. Y., & Little, R. J. A. (2001). Inference for the complier-average causal effect from longitudinal data subject to noncompliance and missing data, with application to a job training assessment for the unemployed. *Journal of the American Statistical Association, 96,* 1232–1244.