Statistical Science
2020, Vol. 35, No. 3, 400–403
https://doi.org/10.1214/20-STS794
Main article: https://doi.org/10.1214/19-STS720

Rejoinder: Linear Mixed Models with Endogenous Covariates: Modeling Sequential Treatment Effects with Application to a Mobile Health Study

Tianchen Qian, Predrag Klasnja and Susan A. Murphy

We thank the Editors for this opportunity and the discussants for their insightful comments. The discussions added much depth and breadth to this work. In the following, we address the comments by each discussant.

1. KRISTIN LINN

1.1 Potential Violation to the Conditional Independence Assumption

The conditional independence assumption in display (10) of the main manuscript says that the time-varying covariate is conditionally independent of the random effects given all the observed history prior to the covariate (i.e., all the previous outcomes, previous treatment assignments and previous covariates). Linn commented that the conditional independence assumption may require further justification for our data example, HeartSteps. In particular, she noted that latent factors, such as levels of depression and anxiety, can be associated with endogenous time-varying covariates such as location, and these latent factors can be part of the random effects. We would like to clarify that this possible association does not necessarily violate the conditional independence assumption; a violation would be such association *after* conditioning on the observed history. For our data example, it is possible that even after conditioning on the observed history, the current location of the user may still be correlated with the latent factors. Because this assumption is not testable without additional modeling assumptions, it is critical to carefully scrutinize the assumption both through scientific knowledge and through sensitivity analyses. Under additional modeling assumptions, it is possible to test for this assumption, and another discussant, Cho et al., provided a potential solution to test empirically the conditional independence assumption (see Section 3 of this rejoinder).

Linn also pointed out that the latent factors, such as levels of depression and anxiety, can change over time and be impacted by prior treatments. For such settings, latent variable models such as partially observed Markov decision processes may prove useful (e.g., Ross et al., 2011).

1.2 Choice Between Marginal Effect and Conditional Effect

Linn commented that a marginal effect estimate may be preferred over a conditional effect, due to reasons including the interest on population-level intervention and ethical and privacy concerns. We agree that there are settings where the marginal effect is of primary interest, such as in the primary analysis of an MRT. For these settings, we recommend use methods for estimating marginal causal effect for MRT (Boruvka et al., 2018, Qian et al., 2019). There are, however, settings where the person-specific effect is of interest. For example, in exploratory analyses behavioral scientists may want to understand how individuals respond to the treatment differently in addition to that explained by the difference in their observed covariates. Such analyses has the potential to aid the development of individualized mobile health interventions (Nahum-Shani et al., 2018).

1.3 Two Practical Suggestions

We wish to echo two other comments by Linn, as they have very important practical implications.

One of our main contributions is to show that under the conditional independence assumption, standard linear mixed model (LMM) software can be used to estimate the conditional effect even if there are endogenous covariates. As a cautionary note, Linn commented that "[e]ase of implementation also opens the door for quick application of the proposed estimation approach by nonstatisticians or time-deprived statisticians who may not take time to fully consider whether the critical assumptions seem reasonable." We reiterate the two caveats when applying standard LMM software with endogenous covariates: (i) When the conditional independence assumption

Tianchen Qian is Postdoctoral Fellow, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: qiantianchen@fas.harvard.edu). Predrag Klasnja is Assistant Professor, School of Information, University of Michigan, Ann Arbor, Michigan 48109, USA (e-mail: klasnja@umich.edu). Susan A. Murphy is Professor, Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, USA (e-mail: samurphy@fas.harvard.edu).