

RE: #PMET-799R1, "Predictive Inference Using Latent Variables with Covariates"

Dear XXX:

We have received the reports on your revised manuscript, "Predictive Inference Using Latent Variables with Covariates" from the Associate Editor (AE) and three reviewers. Reviewer one has some concerns about the paper that were not addressed adequately in the revision. The other two reviewers are largely satisfied now with the paper, although Reviewer three notes that you may benefit from citing and consulting the literature on "congeniality" in multiple imputation. Reviewer one has two major concerns. First, he questions the conditioning arguments presented in Section 4 of the revision because X is not included properly. I think that you need to evaluate his criticisms here and take the necessary steps to resolve them. Second, he argues that in the new real example, there are other potential explanations for the apparent bias induced by choice of model here, and he lists some of them. This point also needs to be addressed. Given that this is a real example and we cannot know the source of differences shown in Table 2, it is important to consider alternatives.

I would like to give you the chance to respond to Reviewer one's points, and so I will reject this version of the paper, and encourage you to revise the paper as needed to address these points. When preparing your revised manuscript, please carefully consider the reviewer comments which are attached, and submit a list of responses to the comments. Your list of responses should be uploaded as a file in addition to your revised manuscript.

We look forward to receiving your revision by . Should you need more time, kindly let us know by return e-mail.

Thank you for your interest in Psychometrika. We look forward to receiving the revision.

Sincerely yours,

XXX

COMMENTS FOR THE AUTHOR:

Reviewer #1: This manuscript is actually now weaker than before.

1) My prior concerns were not addressed, to the contrary the meat of the paper was shortened to pages 18-21, which in addition still contain false assumptions. The conditional distribution of θ used

to argue in section 4 and the theorems and corollaries therein are based on the posterior distribution of theta given Y and Z only (which is wrong). The dependency on item responses X is central here, but completely omitted from the central derivations that 'accuse' the use of PVs as independent as introducing 'wrong model' (a quite subjective term that sets up the reader's expectations rather than carrying any substance) bias. This wrong assumption made in theorem 4.1 and elsewhere in the section of course leads to a model that is completely determined, since all parts are, when $P(\theta, Y, Z)$ is known since then $P(\theta) = F(Y, Z)$ (which is not what is true for 'institutional PVs'. To the contrary, theta is mainly determined by X (the cognitive indicators - the variables most directly associated with theta - this association is much stronger than that between theta and Z, Y in typical cases), i.e. $P(\theta) = G(X, Y, Z)$ (almost = $H(X)$, and the more items in X the more $p(\theta) = H(X)$, note that the NALS example also has a 3 dimensional theta = theta1, theta2, theta3, so the one dimensional theta PV used in the example is probably more determined by the items on the associated scale X1 and other 2 scales X2 and X3 than on any Y, and Z).

2) The simulation study that made some faulty assumptions was not redone with correct assumptions, probably because this would have shown that the far reaching conclusions drawn from this simulation were unsubstantiated.

3) The strong and almost dramatic language used in the first version that talked about how the Goldilocks approach has to be chosen, so that the dependent variable is 'never to be included' in the conditioning model if the PV is used as IV is completely gone from this revision. This is probably also because the initial derivations and logic had issues (as pointed out by prior reviews) that made this Goldilocks requirement unsubstantiated.

4) the empirical example does not discuss potential alternative reasons for the differences found. i) the IRT model used in the example is different from the model used in the 'institutional' analysis. ii) the PVs generated for the NALS were likely based on a multidimensional conditioning model, iii) the MESE implementation needs to be checked against alternative implementation of similar models (using MPLUS, Latent Gold, or so)

5) Finally, the results show that the PV based regression to a large extent agrees with the MESE model with covariates. It seems that the 2 MESE models (with and without covariates) show a much larger WRONG MODEL bias, namely if covariates are omitted we have the so-called omitted variables case, which leads to an overestimate of the race effect since the wrong MESE model (w.o. covariates) fails to include these relevant variables. Interestingly, the race effects are

inflated, but not the effect of the PV skill variable for the wrong MESE model.

Given these issues with the revision, I must conclude in my assessment, that there is not much left that warrants publication or another round of review. In addition, it appears that much of the initial thrust of the paper is mute and not contained in the revision, and hence the papers original points appear to not apply.

===

Reviewer #2: I verified that all comments by the reviewers were taken care of. I suggest to publish the article essentially in its current form.

===

Reviewer #3: I have been out of deep contact with the psychometric world for years, and without doing more work than I have time to do now, the one suggestion that I can make for this paper is to try to build a bridge between the work on this topic with "plausible values" and the more general work on congeniality with multiple imputation. It might even be that some of the specific ideas proposed in the PV literature would have direct extensions to the MI world -- I would think so.

I hope that this brief suggestion is helpful. In general, based on a cursory review of the paper and the reports that you already have, the work appears competent and relevant.

A central reference on the issue of "congeniality with multiple imputation" is provided by:

Meng, X. L. (1994). Multiple-imputation inferences with uncongenial sources of input. Statistical Science, 538-558.
