5.2 Other Considerations in Organization­al Research EPSY 8268

**Heterogeneous Level-1 Variance**

As we estimate a single level-1 within-group variance, Var(*rij*) = σ2, in many applications of HLM for organizational research, we find that this assumption is untenable.

As group comparisons in ANOVA and ANCOVA are somewhat robust to violations of equal variances, HLMs are also relatively robust to violations of the homogeneity of level-1 residual error. But this is sometimes of interest and has implications for model specification.

One possibility is to engage in exploratory research to identify variables that explain level-1 within-group heterogeneity of error variance. This variability in within-group variance might be a function of level-1 or level-2 characteristics. We can model the within-group variance across groups. Since it is a variance, and variances are positive and their sampling distributions are skewed, it is often more helpful to model the log of variance, or ln(σ2), as

ln(σ2) = $α\_{0}+\sum\_{}^{}α\_{0}C\_{j}$.

C*j* can be a level-1 or level-2 predictor. These α coefficients are estimated through maximum likelihood methods and can be tested against a null-hypothesis with a *z*-test, under large-sample theory.

The presence of heterogeneity of level-1 variance can be an indicator of model misspecification. This can be caused by an important omitted variable or if the level-1 explanatory variable was fixed (random variance component set to zero) when it should have been random. Such model misspecification can bias estimates of γ and **T**. Although we can model this heterogeneity (an extra specification in the HLM software), it does not adjust for the problem. To do that, we need to properly specify the model.

**Monitoring Variance Explained**

In the case of random-intercepts-only models, means-as-outcomes models, we are interested in the proportion of variance explained at level 1 and the proportion of variance explained in the level 2 intercepts.

The inclusion of explanatory variables at level 1 will likely reduce the level-1 residual variance, but it may also affect the level-1 variance, τ00. As variables are added to the level-1 model, this changes the meaning of the intercept – with the except of when the explanatory variables are group-mean centered. In this case, β0*j* remains the group mean.

However, because of the complex estimation of variances, the introduction of explanatory variables at level 1 may reduce or increase level-2 variance. That makes the variance explained at level 2 conditional on the level-1 model.

*Step 1 Recommendation*:

Develop the level-1 model first. Evaluate the random variance components for each level-2 residuals, and decide which variables to fix. This then becomes the ***base*** model from which variances are monitored as explanatory variables are added to the level-2 model.

The variance-explained statistics at level 2 will become more complicated when there are multiple random effects at level 2, including the intercepts and multiple slopes as outcomes models. The challenge is based on the extent to which the level-2 residuals are correlated (**T** as correlations). Because of this, when residuals are correlated, entering an explanatory variable into one level-2 slopes-as-outcomes model might affect the variance of another slopes-as-outcomes model. This potentially suggests model misspecification.

If by including an explanatory variable in one level-2 model, the residual variance increases in another model, this may indicate that the explanatory variable belongs in both.

*Step 2 Recommendation*:

In level 2, build the intercept (means-as-outcomes) model first. If additional explanatory variables are under consideration for the slopes models, they should also be added to the intercept model. When random slopes are correlated (at least moderately), any explanatory variable added to one of the slopes models should be added to the others.

When variables are non-significant, they can be removed. But remember that variables that are added to slopes models are interaction terms with the associated variable in level 1 (for which the slopes are estimated). By putting those level-2 explanatory variables in the intercept model, we ensure that the main effects are estimated. Even though these may be nonsignificant, it is probably wise to retain them when the interact effect is significant.

Centering can have odd effects on these complications, especially when a given model uses mixed forms of centering (centering some variables, sometimes with grand-mean or group-mean centering, and not others).

*Step 3 Recommendation*:

For each level-1 explanatory variable (regardless of whether it is centered or how it is centered), include the group mean ($\overbar{X}\_{•j}$) aggregate value of that variable in the intercept model.

This represents the inclusion of both βw and βb in the model, since in most organizational applications, these two values are different. If in fact they do differ, not including the aggregate group mean results in model misspecification.