Adapting Croon's correction for multi-group analyses with equality constraints

As discussed in Devlieger, Mayer, and Rosseel (2016), one can obtain corrected estimation and inferences with factor scores ($\mathbf{Ay}$) by using the corrected covariance matrix, rather than the observed covariance matrix of the factor scores, as input for path analysis. The correction is a function of $\mathbf{A}$ as well as the loading matrix linking the items and the latent variables in the measurement model, as detailed in the appendices of Devlieger et al. (2016). Using the corrected covariance matrix as input is not sufficient for inferences, however, as the analysis ignores that the variables in the input matrix are not observed and contain measurement error.

First consider a simple single-group regression $\eta_Y = b_0 + b_1\eta_X + \zeta$. When both variables are observed, a typical estimator of standard error of the least square estimator of the regression coefficient is given by $\sqrt{SS_\zeta/SS_{\eta_X}/(N-2)}$, where SS_ζ is the error sum of squares and can be estimated by $(N-1)[\mathrm{Var}(\eta_Y) - b_1^2\mathrm{Var}(\eta_X)]$, and SS_{η_X} is the sum of squares for the predictor, and can be estimated by $(N-1)\mathrm{Var}(\eta_X)$. When η_X contains measurement error, however, the error sum of squares is underestimated while the predictor sum of squares is overestimated. To correct for that, one should instead treat the predicted variance of η_Y due to the unreliable variance of η_X as error, and divide by the sum of squares of the factor score predictor:

$$\text{corrected } \hat{SE}(\hat{b}_1) = \sqrt{\frac{\mathrm{Var}(\eta_Y) - \hat{b}_1^2\rho_{\tilde{\eta}_X}\mathrm{Var}(\eta_X)}{(N-2)\rho_{\tilde{\eta}_X}\mathrm{Var}(\eta_X)}},$$

which is equivalent to the standard error given in equation (15) of Devlieger et al. (2016) (p. 749).

For multiple-group analysis, the above standard error estimator can be used separately across groups when there are no equality constraints on the path coefficient. However, when the path coefficient is contrained to be equal, which can be done in standard SEM software such as *lavaan* and *OpenMx* with the group-specific corrected covariance matrices as input, the standard error needs to be "pooled" from the multiple groups. In our simulation Study 1, we used an ad hoc procedure by replacing the $\mathrm{Var}(\eta_Y)$ and $\rho_{\tilde{\eta}_X}\mathrm{Var}(\eta_X)$ term in the above equation with the pooled versions across groups. For example, the pooled $\mathrm{Var}_p(\eta_Y) = \frac{1}{N-G}\sum_g(N_g-1)\mathrm{Var}(\eta_{Yg})$, where N is the total sample size and G is the number of groups. Also we divide by $N-G-1$ instead of $N-2$ to account for the degrees of freedom lost when estimating the group means (i.e., $N-3$ in Study 1 with two groups.).

Devlieger, I., Mayer, A., & Rosseel, Y. (2016). Hypothesis testing using factor score regression: A comparison of four methods. *Educational and Psychological Measurement*, *76*(5), 741–770. https://doi.org/10.1177/0013164415607618