Methods Outline

Jimmy

Methods Outline

Introduction (Three Methods)

Suppose there are two latent predictors (e.g., X and M) and one latent interaction term, XM, in the simplest case. Both the seminal work of Busemeyer and Jones (1983) and Kenny and Judd (1984) provided a classical structural model for estimating latent interaction and nonlinear terms:

$$y(\eta) = \alpha + \gamma_1 \xi_X + \gamma_2 \xi_M + \gamma_3 \xi_X \xi_M + \zeta. \tag{1}$$

In this equation, ξ_X and ξ_M are two latent predictors and their product $\xi_X \xi_M$ form the interaction term, where α represents the intercept. The ζ is the disturbance of regression model with a distribution of $\zeta \sim N(0, \Psi)$ where Ψ is residual variance. The coefficients γ_1 and γ_2 represent the main effects, while γ_3 represents the interaction effect. The dependent variable can be either observed scores y or latent variables η composed of observed items. The measurement models for ξ (s) follow the traditional confirmatory factor analysis (CFA) model:

$$\mathbf{X} = \boldsymbol{\tau}_X + \boldsymbol{\Lambda}_X \boldsymbol{\xi}_X + \boldsymbol{\delta}_X, \tag{2}$$

$$\mathbf{M} = \boldsymbol{\tau}_M + \boldsymbol{\Lambda}_M \boldsymbol{\xi}_M + \boldsymbol{\delta}_M, \tag{3}$$

where $\mathbf{X}$ and $\mathbf{M}$ are vectors of observed exogenous variables, $\boldsymbol{\xi}_X$ and $\boldsymbol{\xi}_M$ are the latent variables represented as scalars, $\boldsymbol{\tau}$ is a vector of constant intercept terms, $\boldsymbol{\Lambda}$ is a matrix of factor loadings, and $\boldsymbol{\delta}$ is a vector of measurement errors at the indicator-level. Moreover, the variance of latent predictor is represented by $Var(\xi)$ and measurement error for the ith indicator is represented by $Var(\delta)$. The relations among measurement errors of

22 all indicators construct a square variance-covariance matrix, Θ_δ, in which the diagonal is

23 filled with elements from the vector $\boldsymbol{\delta}$, and off-diagonal elements are fixed at 0.

24   The assumptions of latent interactions model are concerned with normal distribution

25 and autocorrelation: (1) The terms ξ_X, ξ_X, δ_X, δ_M, ζ are random variables with normal

26 distributions; (2) The terms ζ, δ_X and δ_M are homogeneously distributed with means of 0

27 and uncorrelated with each other, and independently distributed of ξ_X and ξ_M. Note that

28 the correlation between two latent predictors (i.e., $\rho_{\xi_X \xi_M}$) is freely estimated.

29   In the classical latent interaction model, mean-centering is applied as a strategic

30 method to enhance the interpretability of parameter estimates and facilitate accuracy and

31 reduce bias on the estimation of interaction effects. Mean-centering involves subtracting the

32 mean of a variable from its individual values, thereby centering the variable around its

33 average. By mean-centering the constituent variables involved in the interaction, researchers

34 can reduce multicollinearity, making it easier to discern the unique contributions of each

35 variable and their joint impact. This practice aligns with recommendations from seminal

36 works such as Busemeyer and Jones (1983) and Kenny and Judd (1984), who have advocated

37 for mean-centering on mean-centering all indcators of independent latent variables to

38 improve the precision of interaction estimates and to mitigate potential issues arising from

39 the correlation between main effects and interaction terms. Recently a more refined method,

40 Double mean centering (DMC), was proposed to account for non-normally distribution

41 indicators (e.g., categorical data). After calculating product indicators, an additional step of

42 mean-centering again on the product indicators is simply added. By doing so, product

43 indicators with DMC always measure the latent product term and can account for skewed

44 and heterogeneous indicators (Lin et al., 2010).

45   Busemeyer and Jones (1983) proposed that one major advantage of modeling

46 nonlinear effects under the latent variable framework was that it could correct biased

47 estimates of path coefficients and corresponding standard errors by controlling the reliability

of indicators (i.e., the amount of measurement error), where as multiple regression analyses using ordinary least squared estimator assumed error-free measurement on predicting variables. Practically it is very restrictive to hold such an assumption in data analytics. Failure to incorporate measurement error typically leads to substantially erroneous inferences of structural parameters that cannot be remedied by increasing the sample size (Busemeyer & Jones, 1983).

Unconstrained Product Indicator (UPI) [Need a diagram]

Based on the structural model, Kenny and Judd (1984) proposed a product indicator (PI) method to implement latent interaction modeling. The latent interaction term, ξ_{XM} could be indicated by a set of product indicators from ξ_X and ξ_M. For a demonstrative example of measurement model, suppose ξ_X and ξ_M are related to three indicators respectively.

$$
\begin{aligned}
x_1 &= \tau_{X_1} + \lambda_{X_1}\xi_X + \delta_{X_1}; \\
x_2 &= \tau_{X_2} + \lambda_{X_2}\xi_X + \delta_{X_2}; \\
x_3 &= \tau_{X_3} + \lambda_{X_3}\xi_X + \delta_{X_3},
\end{aligned}
\tag{4}
$$

$$
\begin{aligned}
m_1 &= \tau_{M_1} + \lambda_{M_1}\xi_M + \delta_{M_1}; \\
m_2 &= \tau_{M_2} + \lambda_{M_2}\xi_M + \delta_{M_2}; \\
m_3 &= \tau_{M_3} + \lambda_{M_3}\xi_M + \delta_{M_3},
\end{aligned}
\tag{5}
$$

$$x_1m_1 = \tau_{X_1M_1} + \lambda_{X_1M_1}\xi_{XM} + \delta_{X_1M_1};$$

$$x_1m_2 = \tau_{X_1M_2} + \lambda_{X_1M_2}\xi_{XM} + \delta_{X_1M_2};$$

$$...$$ (6)

$$x_3m_3 = \tau_{X_3M_3} + \lambda_{X_3M_3}\xi_{XM} + \delta_{X_3M_3};$$

where ξ_{XM} represents the configuration of the indicators (i.e., x_1m_1, x_1m_2, x_1m_3, x_2m_1, x_2m_2, x_2m_3, x_3m_1, x_3m_2, x_3m_3). There are three observed items (i.e., x_1 to x_3; m_1 to m_3) indicating two latent exogenous variable (i.e., ξ_X and ξ_M) separately. Then nine observed product indicators of all possible configuration of pairs (i.e., from x_1m_1 to x_3m_3) form to indicate the latent interaction term ξ_{XM}. Each observed item has their individual parameter estimation of intercept, factor loading, and residual term.

Kenny and Judd (1984) provided theoretical ground with mathematical derivation and empirical demonstration with a concrete example, but some inherent problems resulted in limited use in psychology research at early stage. Marsh, Wen, and Hau (2004) pointed out that the original PI approach required very complicated nonlinear constraints regarding the SEM model specification, which usually caused statistical issue such as convergence problem and extraordinarily large model. These nonlinear constraints are derived based on the formula of indicators of separate latent variables, and used to specify factor loadings and variances associated with items indicating interaction terms. The final constraints of the interaction term are linear combinations of products of factor loadings, latent varaibles, and error terms (Li et al., 1998). Besides, researchers need to fully understand where and how to specify constraints and select indicators for the interaction term, which was challenging and resource-consuming for those without strong background in relative fields. These obstacles led Marsh, Wen, and Hau (2004) to develop the "UPI" approach that does not require most constraints. Cham et al. (2012) summarized PI approaches and listed two constraints the UPI kept: (1) mean structure of $\boldsymbol{\xi}$s; (2) uncorrelated unique factors of δ_X, δ_M, δ_{XM}. Since

ξ_X and ξ_M are assumed to have expected values of 0, $E(\xi_{XM})$ can be represented by the covariance between ξ_X and ξ_M (i.e., $E[\xi_{XM}] = cov[\xi_X, \xi_M] + E[\xi_X]E[\xi_M]$ where $E[\xi_X] = 0$ and $E[\xi_M = 0]$). Adding the mean structure that could be derived from existing model parameters therefore has no impact on the complexity of model but provides more information of parameter estimation. Nevertheless, it turns out that the mean structure is not necessary if using the DMC strategy, further simplifying the model.

The UPI method, compared to the constrained one, is more robust to the violation of multivariate normality assumptions of latent predictors as it omits to constrain loadings and variances. More importantly, Marsh, Wen, and Hau (2004) demonstrated that the UPI and the constrained PI had parallel performance under the normality assumptions, while the UPI did even better on the simulated non-normal data in which such assumptions were violated. Nevertheless the challenge of choosing indicators was still yet to clear. Foldness and Hagtvet (2014) then conducted a large-scale simulation study in which they studied more than 4,000 possible configurations of indicators within a model with two first-order latent predictors. In general, they concluded that the all-pairs configuration (i.e., using all possible combinations of indicators from stemming latent variables) has the best statistical properties especially for heavily non-normal data (Foldness & Hagtvet, 2014).

In our simulation study, we used the all-paired with double-centering strategy for its simplicity and unbiased parameter estimation on the non-normal data ~~under high non-normal conditions~~ (Cham et al., 2013).

Reliability Adjusted Product Indicator (RAPI) [Need a diagram]

The major difference between the RAPI and the UPI method happens in the measurement model. The RAPI method uses composite scores (e.g., mean score, sum score, etc.) as the single indicator for the latent predictors instead of multiple product indicators:

$$X = \tau_X + 1 \cdot \xi_X + \delta_X;$$

$$M = \tau_X + 1 \cdot \xi_M + \delta_M; \tag{7}$$

$$XM = \tau_{XM} + 1 \cdot \xi_{XM} + \delta_{XM},$$

105   where X and M are the single indicators of their corresponding latent predictors (i.e.,

106   ξ_X and ξ_M), and the latent interaction factor, ξ_{XM}, is indicated by the product of the

107   observed items X and M (i.e., XM). Each observed item has their own estimated intercept

108   and residual term. The observed predictors X and M are one type of double-centered

109   composite scores, for example sum score (i.e, $X = \Sigma X_i$ and $M = \Sigma M_i$) in this case. It

110   should be noted that the factor loadings of two predictors and one interaction term are

111   constrained to 1 receptively due to identification issues for SEM models.

112   A significant feature of the RAPI method is that it can account for measurement

113   errors of observed indicators when the original items that produce the composite scores are

114   unreliable, by placing error-variance constraints in the model. Since the factor loadings are

115   all 1, the reliability of the predictors (e.g., X) can be simplified to:

$$\rho_{XX'} = \frac{Var(\xi_X)}{Var(\xi_X) + Var(\delta_X)}, \tag{8}$$

116   in which $\rho_{XX'}$ is the reliability index and $Var(\xi_X)$ and $Var(\delta_X)$ represent the

117   variance of latent factor and error term. With simple algebra transformation, the error

118   variance can be represented as a function of the reliability (Bollen, 1989):

$$Var(\delta_X) = (1 - \rho_{XX'})Var(X), \tag{9}$$

119   and the variance of the latent variable is:

$$Var(\xi_X) = \rho_{XX'}Var(X), \tag{10}$$

in which $Var(X) = Var(\xi_X) + Var(\delta_X)$.

Under the assumption of independently and identically distributed measurement error, Hsiao, Kwok, and Lai (2018) derived the equations for the error-variance constraint of the interaction term:

$$\begin{aligned}
Var(\delta_{XM}) =& \rho_{XX'}Var(X)(1 - \rho_{MM'}Var(M)+ \\
& \rho_{MM'}Var(M)(1 - \rho_{XX'})Var(X)+ \\
& (1 - \rho_{XX'})Var(X)(1 - \rho_{MM'})Var(M).
\end{aligned} \tag{11}$$

This equation is used to preset the constraint for the error variance of the latent interaction effect in the RAPI method. More derivation details can be found in Appdendix A of Hsiao, Kwok, and Lai (2018).

Two-stage Path Analysis with Interaction (2SPA) [Need a diagram]

The 2SPA was first proposed by Lai and Hsiao (2022), which similarly uses reliability-adjustment technique to account for measurement error in multiple congeneric items. One major improvement of the 2SPA method is that it allows the estimated reliability to be observation-specific and therefore can apply to ordered categorical items that are typically not normally distributed (Lai & Hsiao 2022; Lai et al., 2023).

In the first stage of 2SPA model, researchers obtain factor scores for each observation p on each latent predictor, $\widehat{\xi}_p$, and individual-specific reliability, $\widehat{\rho_p}$. Then, a similar structural model to the RAPI method is applied with reliability-adjusted error constraints.

$$\widehat{F}_{X_p} = \tau_{X_p} + \lambda_{X_p}\xi_{X_p} + \delta_{X_p};$$

$$\widehat{F}_{M_p} = \tau_{M_p} + \lambda_{M_p}\xi_{M_p} + \delta_{M_p}; \quad (12)$$

$$\widehat{F}_{XM_p} = \tau_{XM_p} + \lambda_{XM_p}\xi_{XM_p} + \delta_{XM_p},$$

in which $\widehat{F}_{X_p}$ and $\widehat{F}_{M_p}$ are estimated factors scores for the pth individual from the measurement model, and their product, $\widehat{F}_{XM_p}$, is a single indicator of the latent interaction effect.

For continuous items, the error-variance constraints are $(1 - \rho_{XX'})Var(\widehat{F}_X)$ in which $Var(\widehat{F}_X)$ is the constant variance of factor score for all observations. Similar constraints are applied to the other latent variable $Var(\widehat{F}_M)$. The error variance constraint for the interaction term, $Var(\widehat{FM})$, is define in the same way as RAPI, such as a linear combination of variance and disturbance of two latent variables and reliability. Alternatively speaking, the RAPI method is a special case of 2SPA where the composite scores are used for continuous items (Lai & Hsiao, 2022). As for categorical items with non-normality, the reliability becomes $\rho_{XX'_p} = 1 - \widehat{\sigma}^2_{\widehat{F}_{X_p}}$ where the $\rho_{XX'_p}$ is individual-specific reliability and the $\widehat{\sigma}_{\widehat{F}_{X_p}}$ is the estimated standard error of the factor score for the ith individual. In this case, the factor loading needs to be fixed individually to the corresponding reliability $\rho_{XX'_p}$, and the error variance should be set to $\widehat{\sigma}^2_{\widehat{F}_{X_p}} \cdot \rho_{XX'_p}$.

Method

Design of the Simulation

The design of our simulation study was adapted from Hsiao, Kwok, and Lai (2021). We proposed a new method (i.e., 2SPA) for estimating latent interaction effects and compared it to two other approaches (i.e., UPI, RAPI). The simulated population data were generated from a model with preset parameter values:

$$X_i = \tau_{X_i} + \lambda_{X_i}\xi_X + \delta_{X_i};$$

$$M_i = \tau_{M_i} + \lambda_{M_i}\xi_M + \delta_{M_i}; \tag{13}$$

$$Y_i = \tau_{Y_i} + 0.3\xi_X + 0.3\xi_M + 0.3\xi_{XM} + \zeta_i,$$

where the path coefficients of the structural model with two latent predictors and their interaction term being 0.3. The latent variables ξ_X and ξ_{XM} were simulated from standard normal distribution with means of 0 and variances fixed at 1, each indicated by three items (i.e., ξ_X with $[X_1, X_2, X_3]$; ξ_M with $[M_1, M_2, M_3]$). The X_i and M_i indicators and the dependent variable Y_i were all observed continuous variables with normally distributed errors. The τ_{X_i}, τ_{M_i}, and τ_{Y_i} were their corresponding constant intercepts.

In the measurement model, the λ_{X_i} and λ_{M_i} represent the factor loadings of the ith indicator on ξ_X and ξ_M respectively. Specifically, the first two loadings of the latent predictors were set to fixed values of 1.0 and 0.9 (i.e., $\lambda_{X_1} = \lambda_{M_1} = 1.0$, $\lambda_{X_2} = \lambda_{M_2} = 0.9$), whereas the third loading varies with three levels (i.e., $\lambda_{X_3} = \lambda_{M_3} = [0.75, 0.80, 0.85]$). Then, the third indicator should reflect three levels of loading differences with the first and second indicators (i.e., large difference: [1.0, 0.9, 0.75]; medium difference: [1.0, 0.9, 0.8]; small difference: [1.0, 0.9, 0.85]). We included these three conditions to explore the impact of loading differences on simulation results.

The δ_{X_i} and δ_{M_i} were the residual factors of the ith indicator with variances θ_{X_i} and θ_{M_i} respective, whereas the ζ was the error term for Y with variance Ψ_{Y_i}. All of them followed normal distribution with means of 0. They all have homogeneous distribution with central location at 0 and independent of each other, and further independently distributed of two latent variables ξ_X and ξ_M.

For the UPI method, the ξ_{XM} was indicated by the all-pair configuration of the indicators: $x_1m_1, x_1m_2, x_1m_3, x_2m_1, x_2m_2, x_2m_3, x_3m_1, x_3m_2, x_3m_3$. As for the RAPI and

₁₇₇ 2SPA method, the ξ_{XM} was loaded by a single indicator depending on the procedure of each

₁₇₈ method. Specifically, the indicator of the RAPI method was composite scores of indicators of

₁₇₉ ξ_X and ξ_M respectively, while that of the 2SPA method was pre-computed factor scores

₁₈₀ using corresponding indicators.

₁₈₁   The range of researcher-manipulated sample sizes choices covered in published

₁₈₂ research reports about latent interaction methods spanned from 20 to 5,000 (Chin, Marcolin,

₁₈₃ & Newsted, 2003; Lin et al., 2010; Cham et al., 2012; Marsh et al., 2014), with commonly

₁₈₄ seen choices spanning from 100 to 500. In Hsiao et al. (2018), Hsiao, Kwok, & Lai (2021),

₁₈₅ and Chin, Marcolin, & Newsted (2003), the UPI and RAPI methods reached relatively

₁₈₆ unbiased estimates of the interaction effects at $N \geq 500$, whereas the performance under

₁₈₇ small sample sizes deserved more evaluations and comparisons. Hence we chose N = 250 and

₁₈₈ 500 as our sample size conditions.

₁₈₉   As adapted from the study in Hsiao, Kwok, & Lai (2021), we pre-specified three

₁₉₀ population correlations between latent predictors ($\rho_{\xi_X \xi_M}$): 0, 0.3, 0.6, in which the error

₁₉₁ variance of Y, Ψ_Y, is 0.82, 0.766, 0.712 at $\rho_{\xi_X \xi_M} = 0, 0.3, 0.6$. For the RAPI methods, we

₁₉₂ manipulated the reliability of indicators with three levels: 0.7, 0.8, and 0.9. Varied reliability

₁₉₃ level resulted in three levels of error variance of the exogenous variables. Then, congeneric

₁₉₄ items were formed from manipulating error variances of the three indicators proportionally

₁₉₅ for the corresponding latent variables. We kept the proportions of constrained error

₁₉₆ variances of three indicators from Hsiao, Kwok, & Lai (2021) the same, such that the first

₁₉₇ indicator covered 44% of the total error variances, the second indicator covered 33%, and

₁₉₈ 23% of the last indicator. Hence, the current study has a $2 \times 3 \times 3 \times 3$ design with two

₁₉₉ sample size choices, three factor loading combinations, three correlation levels between latent

₂₀₀ predictors, and three levels of reliability.

²⁰¹ **Evaluation Criteria**

²⁰²   We applied four evaluation criteria to evaluate the performance of the three methods

²⁰³ on accuracy and precision in estimating the interaction effects (i.e., γ_3).

²⁰⁴   The standardized bias and 95% confidence interval (CI) coverage rate were used to

²⁰⁵ evaluate the bias, and measure the accuracy, of the interaction effect estimates. The bias (B)

²⁰⁶ is defined as the average difference between the sample estimator and the preset population

²⁰⁷ parameter:

$$B(\gamma_3) = R^{-1}\Sigma_{r=1}^{R}(\hat{\gamma_{3r}} - \gamma_3), \tag{14}$$

²⁰⁸   where R is the total number of replications that counts from 1 to 2,000. The $\hat{\gamma_{3r}}$

²⁰⁹ denotes the estimated interaction effect in each replication r and the γ_3 denotes the

²¹⁰ population parameter (0.3). To calculate the standardized bias (SB), the standard error of

²¹¹ sample estimated parameter should be obtained from replications and divided by calculated

²¹² bias:

$$SB = \frac{B(\gamma_3)}{SE_{\gamma_3}}, \tag{15}$$

²¹³   where SE_{γ_3} is the standard error of interaction parameter. Collins, Schafer, and Kam

²¹⁴ (2001) suggested that an absolute value of ≤ 0.40 would be considered acceptable for each

²¹⁵ condition.

²¹⁶   As for the 95%CI coverage rate, it was calculated as the proportion of replications in

²¹⁷ which the Wald CI for each replication captures the γ_3. A coverage rate of $> 91\%$ is

²¹⁸ considered acceptable (Muthén & Muthén, 2002).

²¹⁹   Next, the relative standard error (SE) bias was used to evaluate the precision of the

interaction effect estimates. This criterion compared the empirical standard deviation of the interaction estimates with the average standard error across replications:

$$Relative\ SE\ Bias = \frac{R^{-1}\Sigma_{r=1}^{R}(\widehat{SE}_r - SD)}{SD}, \tag{16}$$

where $\widehat{SE}_r$ is the standard error of interaction effect estimate in a single replication cycle and SD is the empirical standard deviation of the interaction estimates obtained from all replications. Smaller relative SE bias represents less variability across replications in each simulation condition, which is more desired. Estimators with absolute values of RSE bias $\leq 10\%$ would be regarded acceptable and indicate that the standard error is reasonably unbiased (Hoogland & Boomsma, 1998).

The last criterion was the root mean squre error (RMSE), calculated by taking the squared root of the sum of squared bias:

$$RMSE = \sqrt{R^{-1}\Sigma_{r=1}^{R}(\hat{\gamma}_{3r} - \gamma_3)^2}. \tag{17}$$

It quantifies both the bias and variability of the interaction effect estimates across replications. A smaller RMSE value would indicate that the replication result of one condition across 2,000 replications had relatively more accuracy and precision.