Moderation
調節 = 交互作用

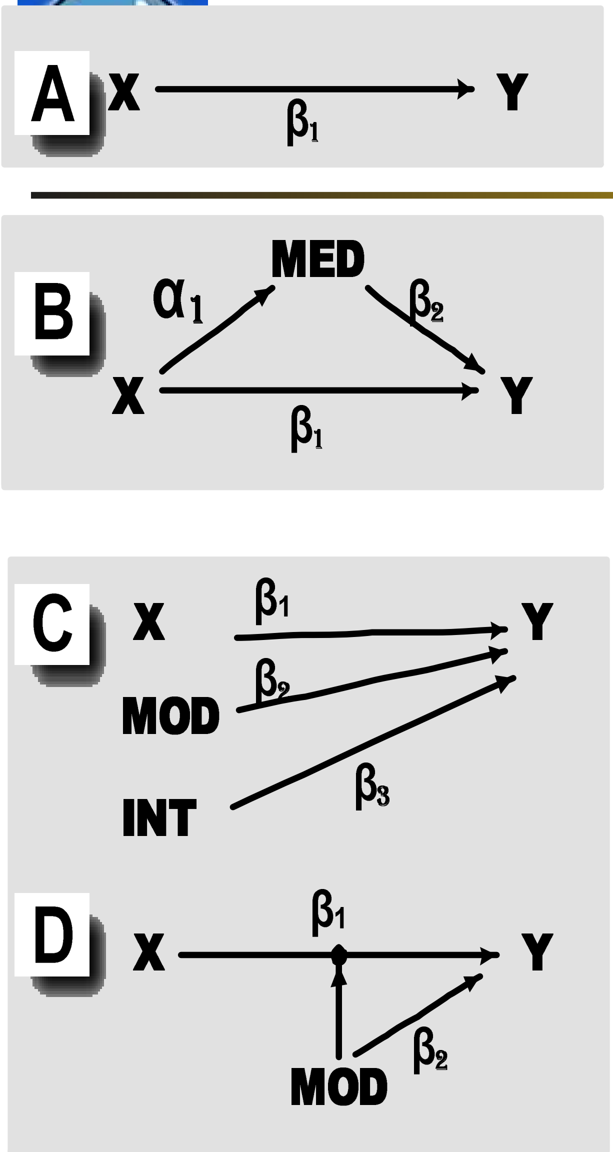
Kit-Tai Hau侯傑泰
JianFang Chang常建芳

The Chinese University of Hong Kong

Based on Marsh, H. W., Hau, K. T., Wen, Z., Nagengast, B., & Morin, A. J. S. (in press). Moderation. In Little, T. D. (Ed.), Oxford Handbook of Quantitative Methods. New York: Oxford University Press.

Introduction

- Examples
  - Ed Psych: effects of an instructional technique interact with students' characteristics
  - Dev Psych: effects of a variable interact with age
  - Soc Psych: effects of individual characteristic depends on Group
  - Organizational Psych: employee characteristics $\times$ workplace characteristics
  - Moderator: variable affects direction and/or strength of relation between indep var and dep var, typically defined as $X_1 \times X_2$
  - (Moderation = interaction) $\neq$ mediation



Mediation (MED) vs Moderation (MOD) X = predictor variable

Y = outcome variable

INT = interaction

A: no mediation or moderation

B: effect of X on Y mediated in part by MED; total mediation if direct effect of X on Y is zero ($\beta 1=0$) or partial medication (if $\beta 1 \neq 0$); indirect effect of X on Y = $\alpha_1 \times \beta_2$

C, D: represent interaction effect; MOD moderates the relation between X and Y

Introduction

- Moderator: variable that affects the direction and/or strength of relationship between independent (predictor) variable and dependent (outcome) variable
- when/for whom does X has a stronger/ weaker relation/effect on Y?

Introduction

Traditional (nonlatent) Approaches

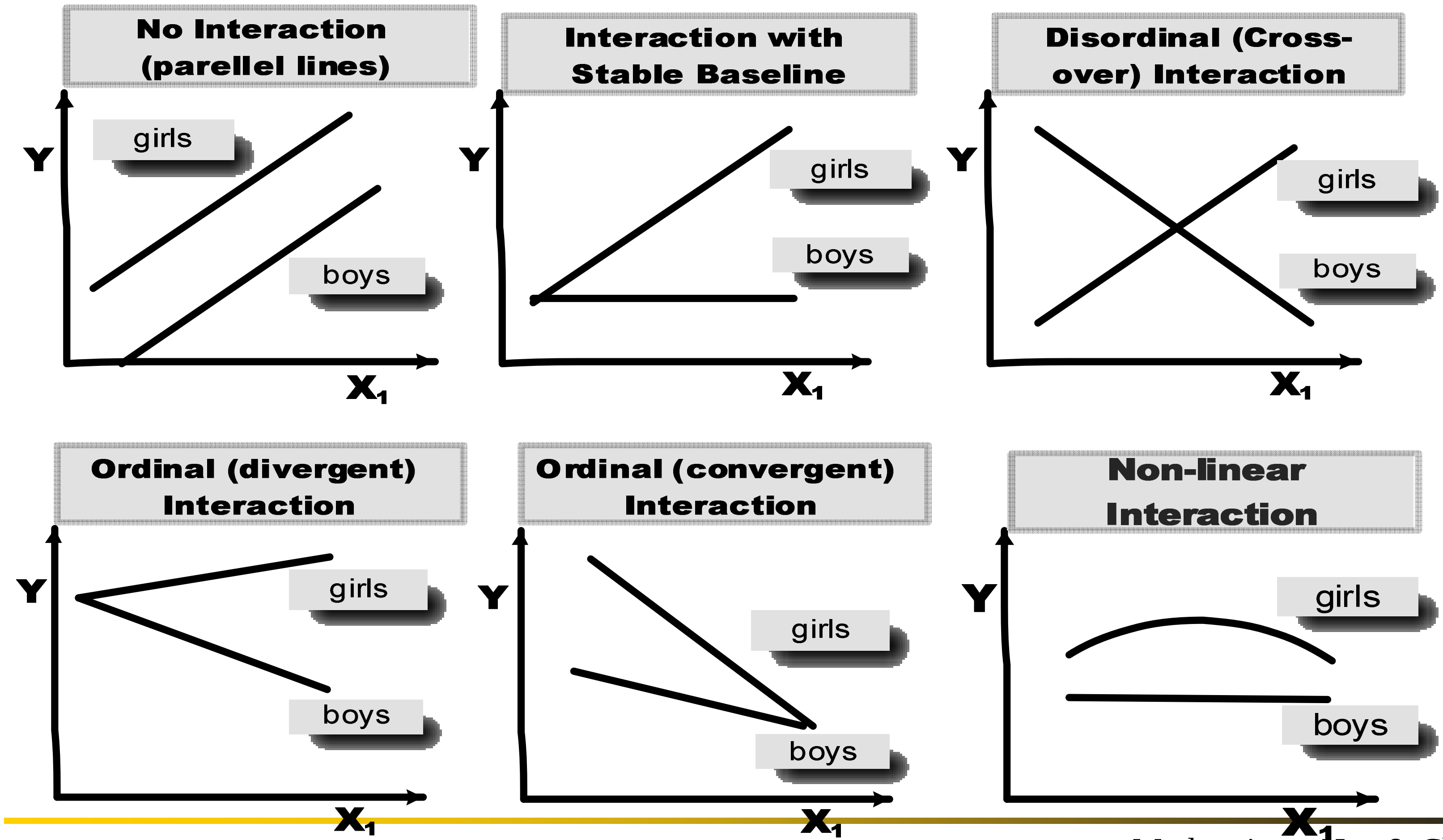
- Interaction between two manifest variables (X_1, X_2) on outcome (Y)
  - X_1, X_2 small number of categories: ANOVA
  - X_1, X_2 cont., regression to estimate main and int'n

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2 + e$$

  - Helpful to graph if interaction is significant
  - Empirical interactions typically small, non-sig, substantial measurement error reduces power of sig test
  - Latent interaction controls for measurement error, increase power, provide more defensible interpretation of interaction

Graphs of Interactions

- inter'n sig ➔ Graph to understand nature



Categorical Variables: Analysis of Variance (ANOVA)

- Null: neither predictor depends on value of other
- tempt to transform continuous variables into ANOVA approach (M split into high/low group)
- Problems: (i) reduce reliability (loss power), (ii) reduce var explained by original variables, (iii) no summary estimate of strength of interaction, (iv) difficult to detect non linear effects
- Transform only when categorization is natural (e.g., minimal passing score)
- Summary: almost never transform cont var -> categories for ANOVA

Separate Group Multiple Regression

- One indep var is categorical (few levels, e.g., gender), another indep var continuous ➔ tempted to use separate group regression
- Inter'n = differences between unstd regression coefficients, possible sig test for 2 groups
- Weakness: not facilitate interpretation of effects (inter'n sig? uncertain); reduce power due to small N in each group; one IV must be truly categorical
- can use the more general approach

Moderated Multiple Regression Approach

- $Y = b_0 + b_1X_1 + b_2X_2 + e$ ➔
- effect of X_2 moderated by X_1 ➔

  $Y = (b_0 + b_1\,X_1) + (b_2 + b_3X_1)\,X_2$
- $Y = b_0 + b_1X_1 + b_2X_2 + b_3X_1X_2 + e$ *standardized* ➔

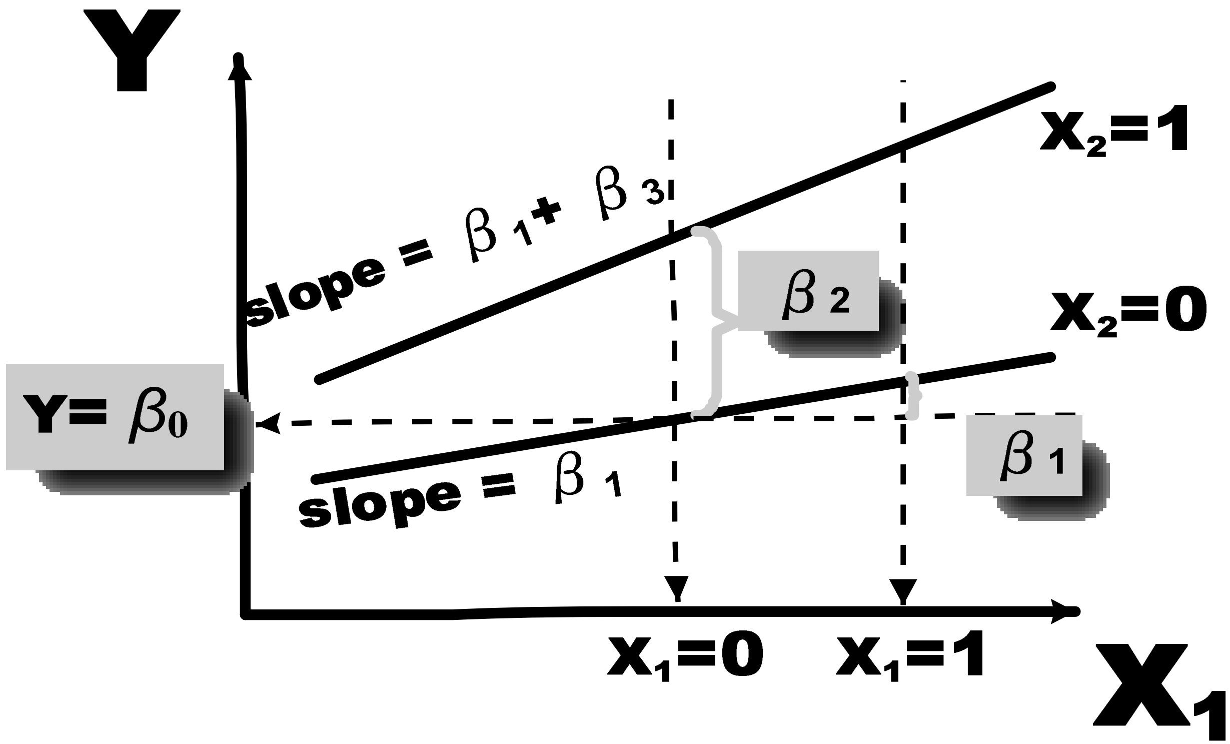
  $Z_Y = \beta_{Z0} + \beta_{Z1}Z_{X1} + \beta_{Z2}Z_{X2} + \beta_{Z3}Z_{X1}Z_{X2} + e_Z$

  β_{z1} = change in Y (in SD unit) if X_1 change 1 *SD* at X_2=0
- Regression plots represent predicted values based on the model (not raw data)
- More than 2 groups -> dummy/effect coding

Moderated Multiple Regression - 2

- Use meaningful zero point; typically at 0, ±1 (or 2) SD

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$$



Moderated Multiple Regression - 3

- Effects of X_1 and X_2 on Y are not unconditional main effect; depends on values of other variable
- X_1, X_2 seldom take 0 (or arbitrary) ➔ regression wt on centered and std variables often more useful
- Std coefficients not straightforward from commercial stat packages; to obtain:
  - Std (z-score) all variables Z_Y, Z_{X1}, Z_{X2}
  - form interaction term = $Z_{X1} \times Z_{X2}$ *(but not re-std)*
  - *Predict* Z_Y, *with* Z_{X1}, Z_{X2}, $Z_{X1}Z_{X2}$ *, use* t-values, and report unstd coef as appropriate std solutions

Moderated Multiple Regression - 4

- regression models testing the interaction term must contain the main effects of predictor variables (X_1,X_2)
- for categorical var involving more then 2 dummy var, all related product terms entered simultaneously, test sig of change in R^2 with/without interaction terms
- covariates (gender, age) generally added as 1st set of variables (but check rationale)
- Regions of significant (range of values of moderator in which simple slopes are significant can be plotted (see Preacher et al. 2006 J Ed Beh Stat 31, 437-448)

Moderated Multiple Regression - 5

- Interaction Point: for disordinal (crossing) interaction, the intersection point for X_1 as the moderator, $X_2 = -b_1/b_3$
- Power in detecting interactions: difficult in finding substantially meaningful, statistically significant interactions because
  - Overall model errors in non-expt studies generally larger than controlled expt
  - Measurement are exaggerated when multiplied to form proudcts

Moderated Multiple Regression - 6

- Magnitude of inter'n constrained in field studies when researchers cannot assign participants to optional levels of predictor variables
- Detect inter'n compromised by nonlinearitie of effects

- To have sufficient power:
  - Large N
  - Similar N across subgroups
  - Avoid restriction of range
  - High reliability scales

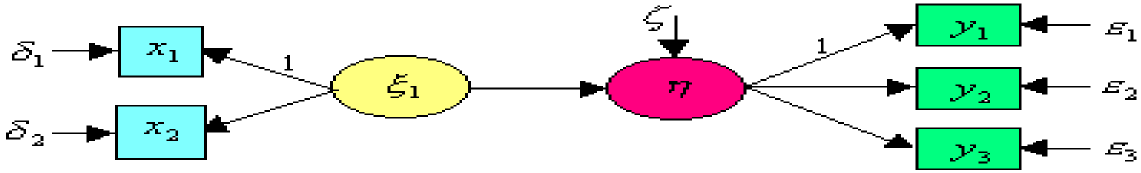
Latent Variable Approaches

- Two Broad categories
  - at least one variable involved is categorical with few categories (e.g., male/female) → multiple group SEM
  - both variables are continuous and latent → various approaches and best practice still evolving

Latent Variable Approach

Multiple Group Analysis

- latent variable (ξ_1) × observed categorical variable (X_2) → latent variable (η)



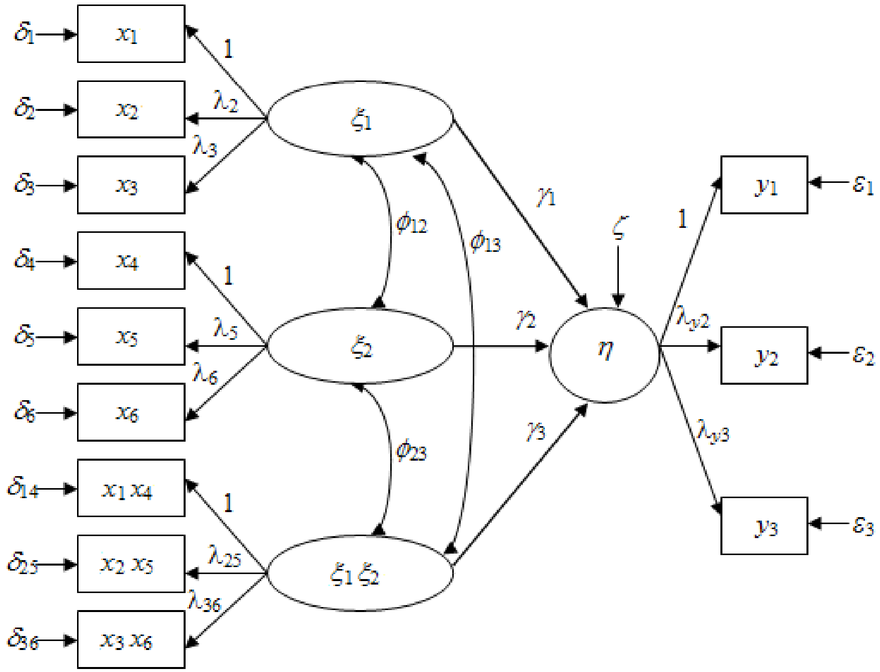
- X_2 small number of naturally existing categories, as grouping var
- test: invariance of ξ_1 → η effects over multiple groups; decline in goodness of fit with invariance constraint
- easily implemented in most SEM software
- problems: limitation in interpretation of the interaction, reduce power (small N), ignore measurement error categorizing var
- Not recommended, unless it is a true catergorial var with small number of categories with at least moderate sample sizes

Latent Variable Approaches

Full Latent (variable) Approach

- Kenny & Judd (1984) proposed an ingenious heuristic model by constraining of loadings/variances of the product term

$$\eta = \gamma_1 \xi_1 + \gamma_2 \xi_2 + \gamma_3 \xi_1 \xi_2 + \zeta$$



Latent Variable Approach

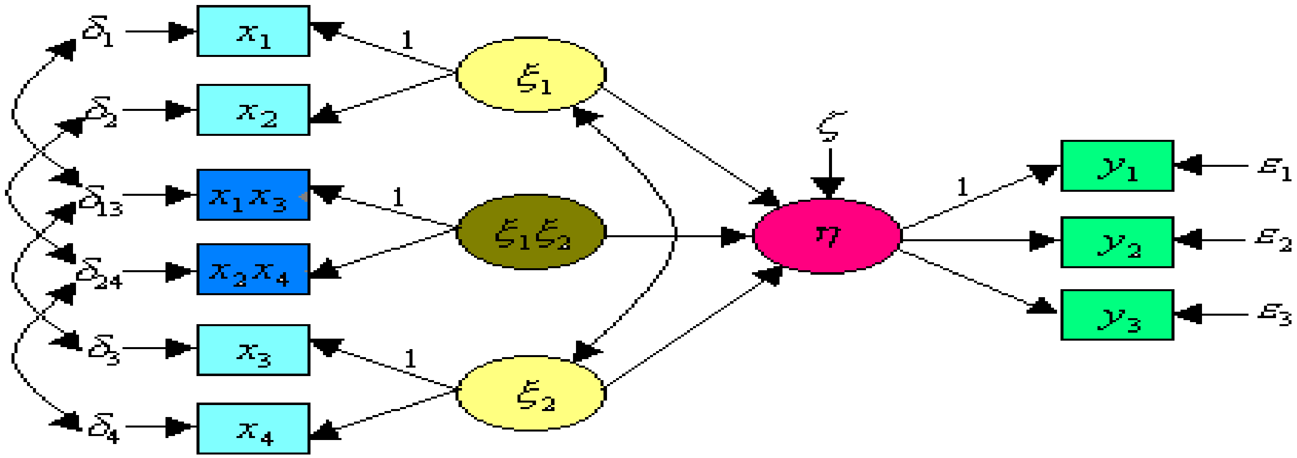
Main Issues

- different ways to form the product indicator; How many product indicators? How to form best set ?
- many constraints on parameters make the method tedious /difficulty, are they absolutely necessary ?
- even if both ξ_1 ξ_2 have mean of zero, product term ξ_1 ξ_2 mean is not zero; mean structure complicates the application, is it really necessary?
- typical software do not provide appropriate SE for std effects, more serious with interaction model, how to obtain appropriate std solution?

Parameter Constrained & UnconstrainedApproaches – Constrained Approach

- Kenny & Judd (1984) proposed an ingenious heuristic model by constraining of loadings/variances of the product term

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3\xi_1\xi_2 + \zeta$$



$$x_1 = \xi_1 + \delta_1 \quad x_2 = \lambda_2\xi_1 + \delta_2 \quad x_3 = \xi_2 + \delta_3 \quad x_4 = \lambda_4\xi_2 + \delta_4$$

Strategies for Creating Product Indicators

- 2 guidelines
  - use all the information (all multiple indicators should be used in forming product indictors)
  - do NOT reuse information: each multiple indicator used once in forming product indicators to avoid artificially created correlated residuals (variance/covariance matrix of errors becomes diagonal)
- Other possible strategies
  - Use the better indicators (when cannot use all indicators)
  - Use parcels (average of indicators) when there are too many indicators in a certain indep var

Parameter Constrained & UnconstrainedApproaches

constrained approach (cont)

- Judd suggested using x_1x_3, x_1x_4,x_2x_3,x_2x_4 as indicators of the interaction $\xi_1\xi_2$ and imposed many constraints on loadings and variances, e.g.

$$x_2x_4 = \lambda_2\lambda_4\xi_1\xi_2 + \lambda_2\xi_1\delta_4 + \lambda_4\xi_2\delta_2 + \delta_2\delta_4$$

- Marsh et al.(2002) simulation showed: unconstrained approach is recommended for its ease in implementation and acceptable bias /precision
- Summary: mean-center all x, y indicators, create product indicator, fit model without mean structure (because software routinely centers them again)

An Appropriate Standardized Solution and Its Scale-free Properties (cont)

- appropriate std solution of interaction model not directly provided by usual commercial software
- Wen, Marsh, Hau (2010) derived appropriate std solution for latent interaction, which are scale free, SE and t-values are also scale free
- Let usual std coefficients be γ_1', γ_2', γ_3', appropriate std coefficients γ_1'', γ_2'', γ_3'' are obtained:

$$\gamma_1'' = \gamma_1' \qquad \gamma_2'' = \gamma_2' \qquad \gamma_3'' = \gamma_3' \frac{\sqrt{\phi_{11}\phi_{22}}}{\sqrt{\phi_{33}}}$$

where $\phi_{11} = \mathrm{var}(\xi_1)$ $\phi_{22} = \mathrm{var}(\xi_2)$ $\phi_{33} = \mathrm{var}(\xi_1\xi_2)$ are from the original solutions

An Appropriate Standardized Solution and Its Scale-free Properties (cont)

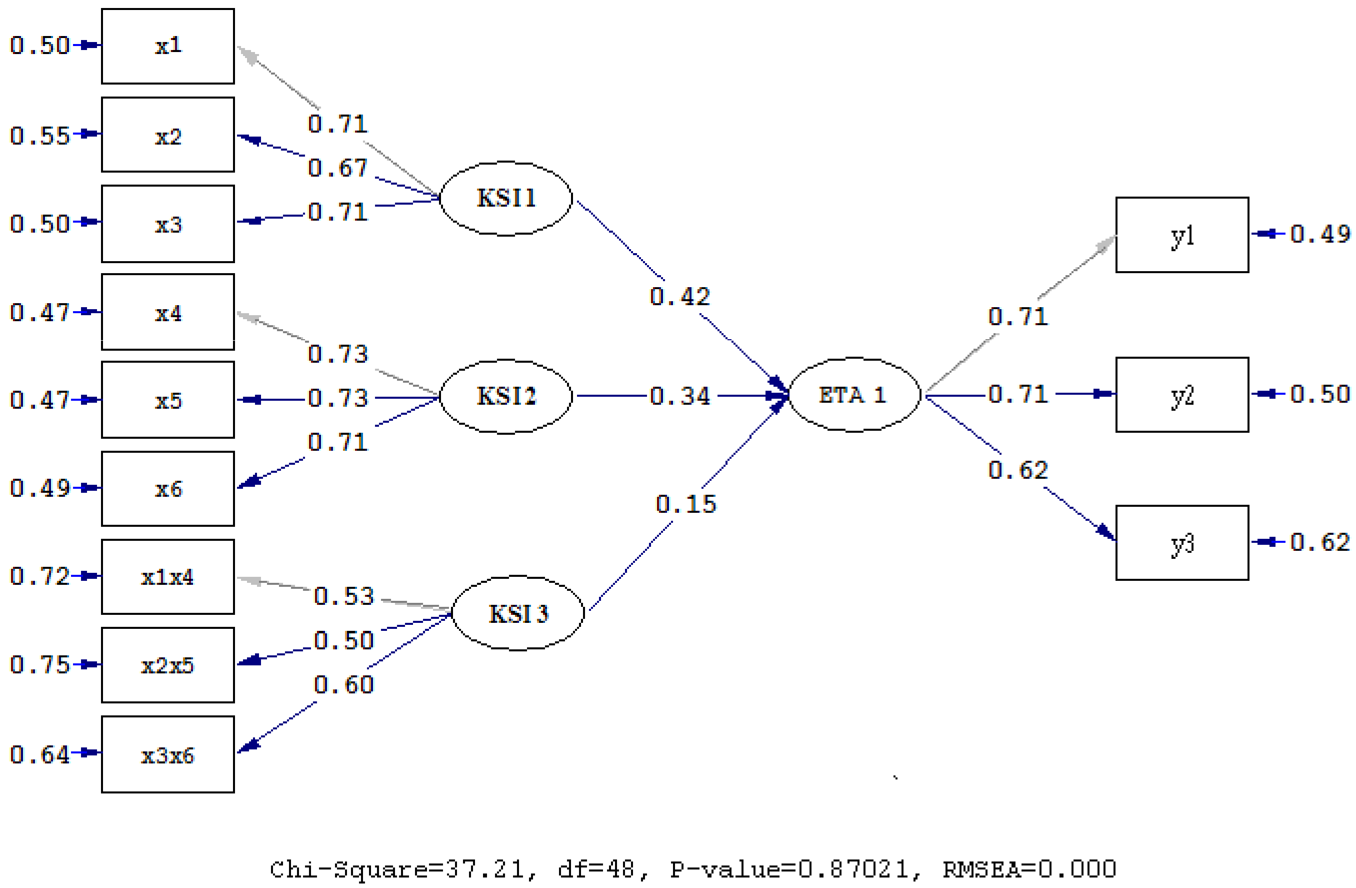
- Scale-free properties of std solution
  - Wen, Marsh, Hau (2010) proved that the appropriate std estimates have the scale-free properties → invariant when calcualted from either the centered or std data
- Calculation of SE of appropriate std coef through Bootstrap samples (similar to original estimates) → t-values of original estimates can be used to test the significance of the appropriate std estimates, if close to cutoff point use bootstrap method

Unconstrained Approaches: Examples

$$\eta = \gamma_1\xi_1 + \gamma_2\xi_2 + \gamma_3[\xi_1\xi_2 - E(\xi_1\xi_2)] + \zeta$$

- Each latent variable has 3 indicators
- Assume η is math achievement, ξ_1 is prior math ability, ξ_2 is math motivation, $\xi_1 \xi_2$ is the interaction of prior math ability and math motivation
- y_1 to y_3, $x_1 \ldots x_6$ centered, product indicators x_1x_4, x_2x_5, x_3x_6 are created, but not re-standardized
- γ_1=0.425, γ_2=0.331, γ_3 =0.197; ϕ_{11}=0.501, ϕ_{22}=0.529, ϕ_{33} =0.308; and the completely standardized estimates: γ_1' =0.423, γ_2' =0.338 and γ_3' =0.153. By using Formula 27, γ_1'' =0.423, γ_2'' =0.338, and $\gamma_3''=0.142$

Unconstrained Approaches: Examples (cont)



el

Robustness to Normality in Unconstrained Approach

- Considerations when normality is violated:
  - ML typically used is based on assumption of normality, however, this is a common problem to all CFA, SEM research (not specific to interaction/quadratic analyses)
  - even when ξ_1,ξ_2 are normal, the product are non-normal, constrained, partially constrained, unconstrained all suffer when ML estimation is used
  - Fortunately, ML tends to be robust to violation of normality in parameter estimates, though ML likelihood ratio test is too large, standard errors are too small under nonnormality

Distribution-analytic Approaches

- Whereas they have many desirable features, they are computationally demanding, and not available in widely accessible SEM softwares
  - Latent Moderated Structural Equation (LMS, Klein & Moosbrugger, 2000) implemented in Mplus
  - Quasi-Maximum Likelihood(QML, Klein & Muthén, 2002) – available from author, not available in software yet
- QML (Klein & Muthen, 2002) was developed for more efficient estimation than LMS
- Both estimate parameters in $\eta = \alpha + \Gamma\,\xi + \xi'\Omega\,\xi + \zeta$
- LMS and QML differ in the distributional assumptions about the latent dependent variable η and its indicators
- Computationally LMS is more efficient and can be used to fit models with a larger number of nonlinear effects and interactions

Summary

- One of predictor variables is a manifest grouping variable with small number of categories → multiple group SEM, but not recommended when all predictors are continuous or based on multiple indicators
- Product-indicator dominated latent interaction research, still evolving, unconstrained approach – ease of implementation and robustness
- More recently, LMS/QML hold considerable promise over product-indicator approach
- Many issues not appropriately dealt with and applied research is limited

Thank You