Chapter 11

Multiple-Sample SEM

Facts do not cease to exist
because they are ignored.

—Aldous Huxley

Overview

- Rationale of multiple-sample SEM
- Multiple-sample path analysis
- Multiple-sample CFA
- Extensions
- MIMIC models as an alternative

Rationale of multiple-sample SEM

- Just about any type of structural equation model can be analyzed across multiple samples

- The main question addressed in a multiple-sample SEM is:

  Do values of model parameters vary across groups?

- Another way to express the same question is:

  Does group membership moderate the relations specified in the model (i.e., is there an interaction effect)?

Rationale of multiple-sample SEM

- Perhaps the simplest way to conduct a multiple-sample analysis is to estimate the same model within each of two or more different groups
- Next one would compare the *unstandardized* solutions across the groups
- Recall that unstandardized instead of standardized estimates should generally be compared when the groups differ in their variabilities

Rationale of multiple-sample SEM

- For the same reason

  1. covariance matrices (or the raw scores) for each group should be analyzed when the model has just a covariance structure

  2. covariance matrices and means (or the raw scores) should be analyzed when the model has both a covariance structure and a mean structure

- If the unstandardized estimates for the same parameter are appreciably different, then the populations from which the groups were sampled may not be equal on that parameter

Rationale of multiple-sample SEM

- More sophisticated comparisons are available by using a SEM computer program that performs a multiple-sample analysis, which simultaneously estimates a model across all samples

- Through the specification of *cross-group equality constraints*, group differences on any individual parameter or set of parameters can be tested

- A cross-group equality constraint forces the computer to derive equal unstandardized estimates of that parameter in all samples

Rationale of multiple-sample SEM

- The fit of the model with parameters constrained to be equal across the groups can then be compared with that of the unrestricted model without the equality constraints with the chi-square difference (χ^2_D) statistic

- If the fit of the constrained model is much worse than that of the unconstrained model, one concludes that the parameters may not be equal in the populations from which the samples were drawn

Rationale of multiple-sample SEM

- The comparison just described assumes that the unconstrained model fits the data reasonably well—otherwise, it makes little sense to impose constraints on a model with poor fit in the first place
- *However, do not forget that estimates constrained to be equal in the unstandardized solution will typically be unequal in the standardized solution if the groups have different variabilities*
- In general, standardized estimates should be directly compared only across different variables within each sample

Multiple-sample path analysis

- The basic rationale of a multiple-sample path analysis (PA) is the same whether the model is recursive or nonrecursive
- The main focus of a multiple-sample PA is often the comparison of estimates of unstandardized direct effects across the groups
- This comparison determines whether the magnitudes or directions of estimated direct effects are appreciably different across the groups
- The hypothesis of equal population direct effects is represented by the imposition of cross-group equality constraints on the estimates of these effects

Multiple-sample path analysis

- The relative fits of path model with equality-constrained direct effects with that of the unconstrained model can be tested with χ^2_D

- If the χ^2_D statistic is statistically significant, then freely estimate the direct effects within each sample and compare these results

Multiple-sample path analysis

- If group differences in values of the unstandardized path coefficients are meaningful, then the hypothesis of equal direct effects may be rejected

- It is also theoretically possible to constrain other parameter estimates to be equal (e.g., the disturbance variances), but this is rarely done in practice

Multiple-sample CFA

- Considered first is the evaluation of CFA measurement models without a mean structure across multiple samples
- The main question of a multiple-sample CFA concerns *measurement invariance*, which is whether a set of indicators assesses the same constructs in different groups
- A related concept in the psychometric literature is that of *construct bias*, which implies that a test measures something different in one group than in another
- If so, then group membership moderates the relation between the indicators and factors specified in the measurement model

Multiple-sample CFA

- The evaluation of measurement invariance with CFA typically involves the comparison of the relative fits with the χ^2_D statistic of two hierarchical factor models
- One of these models is estimated with cross-group equality constraints imposed on some of its parameters and the other is estimated without constraints

Multiple-sample CFA

- A common practice is to constrain the unstandardized factor loadings to be equal across the samples
- If the fit of a CFA model with equality-constrained loadings is not appreciably worse than that of the unconstrained model, then the indicators may measure the factors in comparable ways in each group

Multiple-sample CFA

- If the fit of the constrained model is considerably worse, however, then individual unstandardized factor loadings should be compared across the samples to determine the extent of *partial measurement invariance*

- This means that some factor loadings may vary appreciably across groups but the values of others do not

Multiple-sample CFA

- Although it is theoretically possible to do so, cross-group equality constraints are usually not also imposed on estimates of variances or covariances in a multiple-sample CFA

- This is because groups may be expected to differ in their variabilities on either the common factors or measurement errors even if the indicators measure the same factors in all groups (MacCallum & Tucker, 1991)

Multiple-sample CFA

- As in a single-sample CFA, the factors must be assigned a scale in order for the model to be identified
- However, there are some special considerations when scaling factors in multiple-sample CFA
- It is generally *inappropriate* to scale the factors by standardizing them (fixing their variances to 1.0) in all samples
- Doing so assumes that the groups are equally variable on the factors
- If group variances on the underlying factors are really different, then this method for scaling factors may lead to inaccurate results

Multiple-sample CFA

- A better way to scale the factors in a multiple-sample CFA is to fix the unstandardized loading of one indicator per factor to 1.0
- That is, analyze unstandardized factors
- However, note that the loadings of the same indicator should be fixed to 1.0 in each sample
- That is, each factor should have the same reference variable across all groups
- This method allows the variances of the factors to be freely estimated in each sample

Multiple-sample CFA

- There are two potential complications to the method just described for scaling factors in a multiple-sample CFA:

  1. Factor loadings fixed to 1.0 in all samples cannot be tested for group differences because they are constants, not variables (i.e., they have no standard errors)

  2. Because fixed loadings are excluded from tests of measurement invariance, it must be assumed a priori that the associated indicator assesses the factor equally well in all groups

Multiple-sample CFA

- The potential complications just listed imply that if the researcher decides to fix the loading of an indicator that is not invariant across the groups, then the subsequent results may be inaccurate
- One way to deal with this problem is to reanalyze the model after fixing the loading of other indicators to 1.0
- If the unstandardized factor loadings that were originally fixed are comparable in new analyses in which they are free parameters, then that indicator may be measurement invariant
- See Reise, Widaman, and Pugh (1993) for more information

Multiple-sample CFA

- It is also possible to analyze a CFA measurement model with a mean structure across multiple samples
- This allows hypotheses about both measurement invariance and group mean differences on latent variables to be tested in the same analysis
- The input data for this kind of analysis is a covariance matrix and means (or the raw data) for each group

Multiple-sample CFA

- Recall that a mean structure usually includes direct effects of a constant that equals 1.0 for every case on the indicators or factors

- Because the indicators in a CFA model are endogenous, direct effects of the constant on them are intercepts, but the total effects are means

Multiple-sample CFA

- In contrast, direct effects of the constant on the factors, which are exogenous, are means
- However, also recall that it may be impossible to estimate both indicator intercepts and factor means when a CFA model with a mean structure is analyzed in a single sample
- This is because the full mean structure may not be identified in this kind of analysis (chap. 10)

Multiple-sample CFA

- There is a two-part strategy by Sörbom (1974) to identify mean structures of CFA models analyzed across multiple samples
- It involves the estimation of *relative differences* in factor means instead of *absolute differences*
- The first part of this strategy is to fix the means of all factors to zero in one group

Multiple-sample CFA

- This is the same as constraining the direct effect of the constant on the factors to zero in that group
- These constraints establish that group as a *reference sample*
- The factor means are then freely estimated in the other groups, and their values are relative differences on the factor means

Multiple-sample CFA

- When there are only two samples, it is arbitrary which group is selected as the reference sample
- The choice is more critical with three or more samples because all factor mean differences are estimated relative to the same reference sample

Multiple-sample CFA

- The second part of Sörbom's (1974) strategy concerns the covariance structure, which in this case is a measurement model
- Specifically, in order to reasonably estimate relative group differences on either means or intercepts among the latent variables, it must be assumed that the factors are defined the same way for all samples
- Scaling the factors by fixing the loading of the same indicators to 1.0 in each sample is one way to address this issue
- Another is to constrain both the factor loadings and intercepts of the indicators to be equal across the groups

Multiple-sample CFA

- Constraining both factor loadings and intercepts to be equal across the groups provides a test of whether the equations for the regressions of the indicators on the factors are comparable in each group

- At minimum, the indicators should have the same basic factor structure and reasonably similar unstandardized factor loadings and intercepts across the groups in order to estimate relative factor mean differences

Multiple-sample CFA

- If the hypothesis of invariance does not hold for a couple of indicators, then their factors loadings or intercepts can be estimated separately within each group (i.e., release the equality constraints on those loadings)
- This controls for unequal directs effects of the factors on those indicators
- Otherwise, it makes little sense to estimate relative factor mean or intercept differences if the indicators do not seem to measure the same basic constructs in each group
- See See R. Williams and Thomson (1986) for more information

Extensions

- The multiple-sample analysis of a structural regression model with both exogenous and endogenous factors and a mean structure that is not a latent growth model (LGM) follows the same basic rationale

- For example:

    1. Factor measurement should be specified the same way for all groups

    2. For the group selected as the reference sample, all direct effects of the constant on the factors are fixed to zero in order to identify the mean structure

- The only difference is that direct effects of the constant on endogenous factors are interpreted as relative group differences in the intercepts for the regression of those factors on other variables specified as direct causes

Extensions

- Because of different identification requirements for a LGM, it is generally possible to estimate the means or intercepts of latent growth factors in a single sample

- This implies that latent growth factor means or intercepts can be estimated separately for each group when a LGM is simultaneously analyzed across multiple samples

- These capabilities provide great flexibility in hypothesis testing with longitudinal data (e.g., T. Duncan, S. Duncan, Strycker, Li, & Alpert, 1999)

MIMIC models as an alternative

- An alternative way to estimate group differences on latent variables is through the specification of a MIMIC (multiple indicators and multiple causes) model

- In such a model, factors with effect indicators are regressed on one or more dichotomous cause indicators that represent group membership

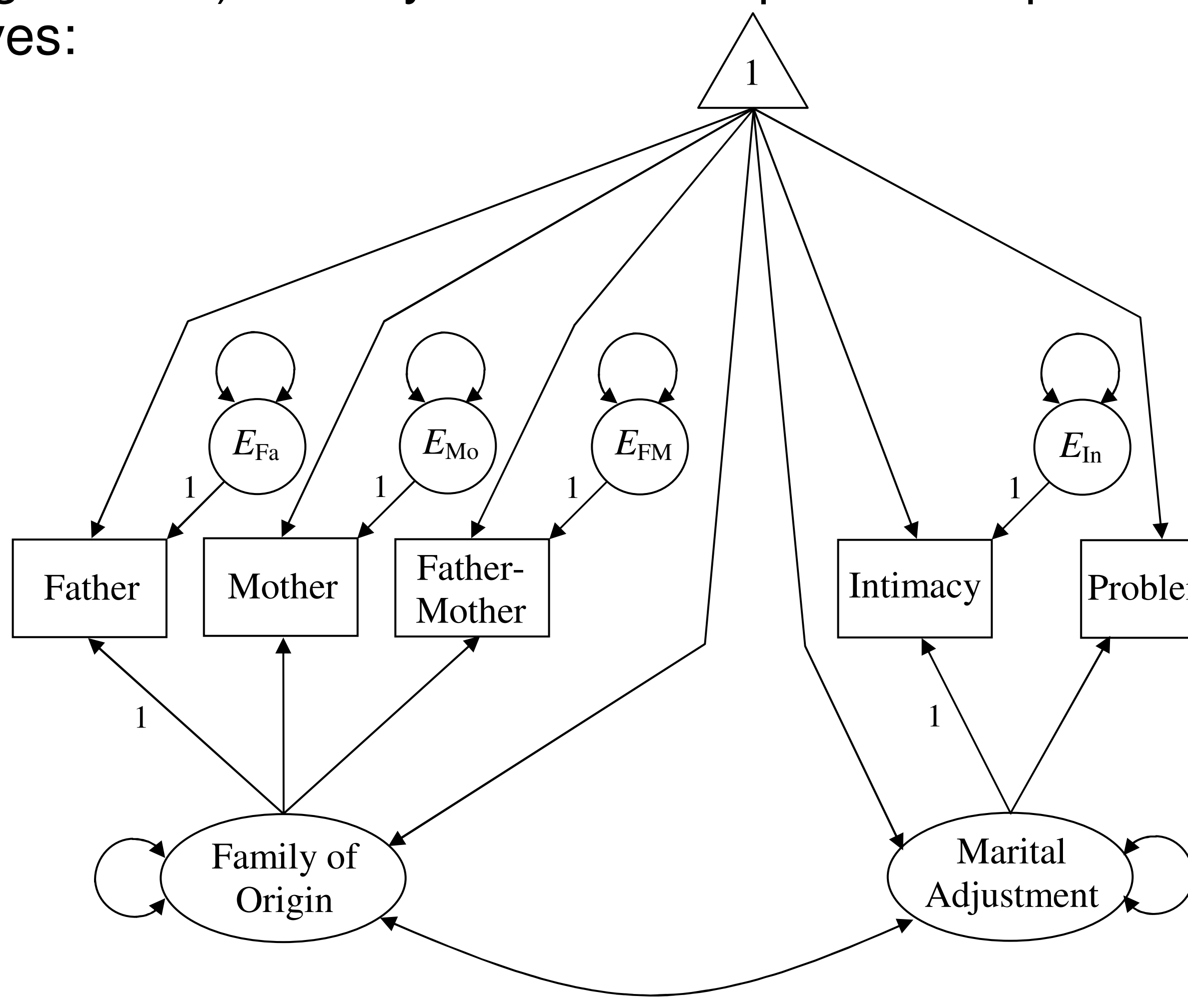
MIMIC models as an alternative

- In this approach, the total sample is not partitioned into subsamples as is the case in a multiple-sample SEM (although subsamples are still required in the study design)

- Thus, there are no special identification requirements beyond the usual ones for single-sample analyses for the types of MIMIC models described here

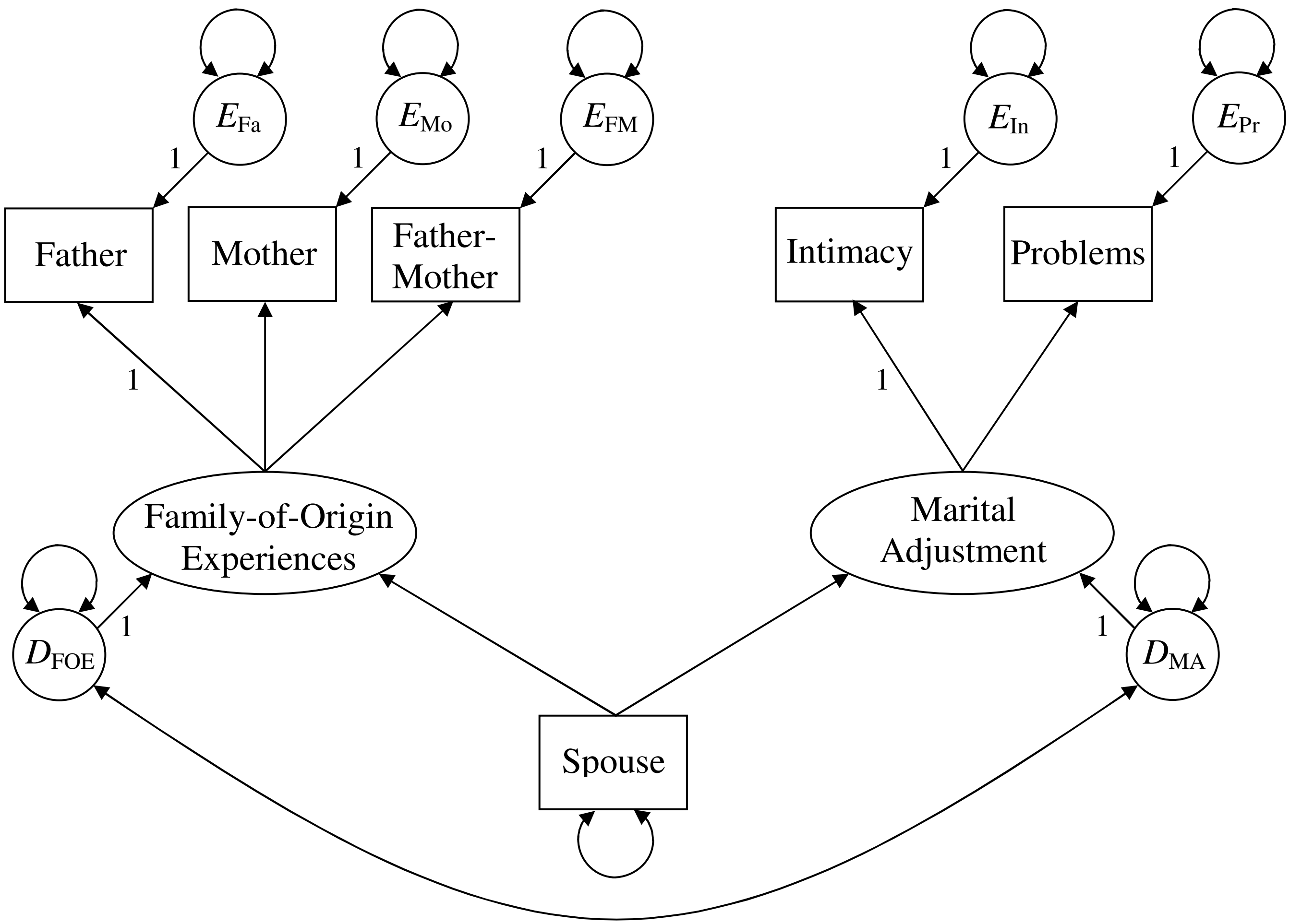
MIMIC models as an alternative

- Suppose that this CFA measurement model with a mean structure (Figure 11.2) is analyzed across separate samples of husbands and wives:



MIMIC models as an alternative

- A MIMIC alternative is presented here (Figure 11.3) that is analyzed with data from the combined sample (husbands and wives together):



MIMIC models as an alternative

- In the MIMIC model just presented:

    1. The single cause indicator in the MIMIC model of Figure 11.3 is a dichotomy that represents spouse coded as 0 = husband and 1 = wife

    2. It is specified to have direct effects on a family-of-origin experiences factor with three effect indicators and a marital adjustment factor with two effect indicators

MIMIC models as an alternative

- In the MIMIC model just presented:

    3. The factors are endogenous in the MIMIC model, and the disturbances are specified as correlated

    4. Although there is no mean structure, the path coefficients for the direct effects of spouse will provide information about the degree to which the difference between husbands and wives predicts the factors

- See Kano (2001) and Kaplan (2000, chap. 4) for additional examples

References

Duncan, T. E., Duncan, S. C., Strycker, L. A., Li, F., & Alpert, A. (1999). *An introduction to latent variable growth curve modeling: Concepts, issues, and applications.* Mahwah, NJ: Erlbaum.

Kano, Y. (2001). Structural equation modeling with experimental data. In R. Cudeck, S. du Toit, and D. Sörbom (Eds.), *Structural equation modeling: Present and future* (pp. 381-402). Lincolnwood, IL: Scientific Software International.

Kaplan, D. (2000). *Structural equation modeling*. Thousand Oaks, CA: Sage.

MacCallum, R. C., & Tucker, L. R. (1991). Representing sources of error in common factor analysis: Implications for theory and practice. *Psychological Bulletin*, *109*, 501-511.

Reise, S. P., Widaman, K. F., & Pugh, R. H. (1993). Confirmatory factor analysis and item response theory: Two approaches for exploring measurement invariance. *Psychological Bulletin*, *114*, 552-566.

Williams, R, & Thomson, E. (1986). Normalization issues in latent variable modeling. *Sociological Methods and Research*, *15*, 24-43.