Meta-analysis: A Structural Equation Modeling Perspective

Mike W.-L. Cheung
Last update: 20 April 2009

Meta-analysis and structural equation modeling (SEM) are two popular statistical techniques in the social, behavioral, and medical sciences. Meta-analysis is used to synthesize effect sizes from a pool of empirical studies whereas SEM is used to fit hypothetical models on primary studies. These two techniques are generally treated as two unrelated areas in the literature. This note provides a brief introduction on how to conduct meta-analysis from a SEM perspective. It is hoped that more research will be devoted to integrating meta-analysis and SEM.

Two types of models are introduced here: SEM-based meta-analysis (Cheung, 2008) and meta-analytic structural equation modeling (MASEM; Cheung & Chan, 2005a, 2009). SEM-based meta-analysis is used to conduct conventional fixed-, random-, and mixed-effects meta-analysis by treating studies in a meta-analysis as subjects in a structural equation model. MASEM is used to pool correlation (or covariance) matrices and to fit structural equation models on the pooled correlation (or covariance) matrix. A multiple-group SEM approach may also be used to conduct fixed-effects meta-analysis (Cheung, in press). To limit the scope of this note, this approach is not addressed here.

SEM-based Meta-analysis

Meta-analytic Models

Fixed-effects models. In this note, y_i represents a generic effect size in the *i*th study, such as raw mean difference, standardized mean difference, log odds ratio, log relative risk, and correlation coefficient and its Fisher's *z* transformed score. y_i is usually expressed as

$$y_i = \beta_F + e_i$$

where β_F and e_i are the population effect size and the sampling error in the *i*th study, respectively. e_i is assumed to be normally distributed with a mean of zero and a known variance of σ_i^2 .

The estimated population effect size $\hat{\beta}_F$ under a fixed-effects model is

$$\hat{\beta}_F = \frac{\sum_{i=1}^{k} w_i y_i}{\sum_{i=1}^{k} w_i}$$

where $w_i = 1/\sigma_i^2$ is the weight and *k* is the number of studies. The estimated sampling variance s_F^2 of $\hat{\beta}_F$ is computed by

$$s_F^2 = 1/\sum_{i=1}^{k} w_i \quad .$$

After obtained the fixed-effects estimate, it is of interest to test whether the estimated population effect size is statistically significant. We may compute a test statistic

$$Z_1 = \hat{\beta}_F / s_F \quad .$$

Under the null hypothesis $H_0: \beta_F = 0$, the test statistic Z_1 has an approximate standard normal distribution.

Random-effects Models. Fixed-effects models assume that the population effect sizes share a common value. Many researchers argue that studies are not direct replications of each other. It is expected that there will be differences in the population effect sizes due to differences with the samples and methods used across studies. Thus, random-effects models should be more appropriate (e.g., Hedges & Vevea, 1998; Hunter & Schmidt, 2000; National Research Council, 1992).

Besides the sampling error, random-effects models include variations in the population effect sizes. The random-effect model is

$$y_i = \beta_R + u_i + e_i \quad ,$$

where β_R , u_i and e_i are the mean population effect size, the study specific effect, and the sampling error in the *i*th study, respectively. In fixed-effects models, there is only one source of variation, the sampling variance σ_i^2 . In contrast, there are two sources of variation in random-effects models – the sampling variance and the between-study variance component, $\tau^2 = var(u_i)$.

One common estimator of τ^2 was proposed by DerSimonian and Laird (1986). Their estimator is

$$\hat{\tau}^2_{DL} = \max\left(0, \quad \frac{Q - (k-1)}{c}\right)$$

where Q is the statistic of the homogeneity test, k stands for the number of studies, and $c = \sum_{i=1}^{k} w_i - \left(\sum_{i=1}^{k} w_i^2\right) / \left(\sum_{i=1}^{k} w_i\right)$. Maximum likelihood (ML) and restricted maximum likelihood (REML) estimations may also be used to estimate τ^2 (see Viechtbauer, 2005).

Once the variance component τ^2 is estimated, the estimated mean population effect size $\hat{\beta}_R$ under the random-effects model is

$$\hat{\beta}_R = \frac{\sum_{i=1}^{k} \tilde{w}_i y_i}{\sum_{i=1}^{k} \tilde{w}_i}$$

where $\tilde{w}_i = 1/(\sigma_i^2 + \hat{\tau}^2)$ is the new weight. The estimated sampling variance s_R^2 of $\hat{\beta}_R$ is computed by

$$s_R^2 = 1 / \sum_{i=1}^{k} \tilde{w}_i \quad .$$

Under the null hypothesis $H_0: \beta_R = 0$, the test statistic $Z_2 = \hat{\beta}_R / s_R$ has an approximate standard normal distribution.

Mixed-effects models. It is sometimes of interest to include study-specific covariates to explain population heterogeneity. These are generally known as mixed-effects models, and are also widely known as meta-regression in medical research (Berkey et al., 1995; Thompson & Higgins, 2002). Mixed-effects models include both fixed- and random-effects. The fixed-effects are the regression coefficients due to the study-specific covariates, while the random-effects are the unexplained study-specific effects after controlling for the covariates.

The model in matrix notation is

$$\mathbf{y} = X\boldsymbol{\beta} + I_k\mathbf{u} + \mathbf{e}$$

where $\mathbf{y}$ is a $k \times 1$ vector of effect sizes, $\boldsymbol{\beta}$ is a $p \times 1$ vector of fixed-effects regression coefficients including the intercept, $\mathbf{u}$ is a $k \times 1$ vector of study-specific random effects with $\mathbf{u} \sim N(\mathbf{0}, I_k\tau^2)$, $\mathbf{e}$ is a $k \times 1$ vector of residuals, X is a $k \times p$ design matrix that includes ones in the first column, and I_k is a $k \times k$ identity matrix. Since the effect sizes are assumed to be independent, the conditional covariance matrix of the residuals $V_\mathbf{e}$ is a diagonal matrix, that is, $V_e = diag[\sigma_1^2, \sigma_{2,}^2 ..., \sigma_k^2]$.

Raudenbush (1994) proposed a method of moments estimator on τ^2 under the mixed-effects meta-analysis. Besides using the method of moments, multilevel models may also be used to analyze random- and mixed-effects meta-analyses (e.g., Hox, 2002; Konstantopoulos & Hedges, 2004; Raudenbush, 1994; Raudenbush & Bryk, 2002). When $\hat{\tau}^2$ is available, weighted least squares can be used to obtain the parameter estimates and their asymptotic covariance matrix by using a new weight $\tilde{w}_i = 1/(\sigma_i^2 + \hat{\tau}^2)$.

An SEM approach

Fixed-effects models. One major issue of using SEM to analyze the meta-analytic data is that the effect sizes are distributed with known variances. This violates the basic assumption in SEM in which the data are distributed with the same variance. To make the effect sizes suitable for SEM, we transform all variables including the intercept by

$$W^{1/2} = \text{diag}[1/\sigma_1, 1/\sigma_2, \ldots, 1/\sigma_k]$$

(e.g., Kalaian & Raudenbush, 1996; Konstantopoulos, 2008; Raudenbush, Becker, & Kalaian, 1988).

After the transformation, the fixed-effects model becomes

$$W^{1/2}\mathbf{y} = W^{1/2}X\boldsymbol{\beta} + W^{1/2}\mathbf{e}$$

$$\mathbf{y}^* = X^*\boldsymbol{\beta} + \mathbf{e}^*,$$

where $\mathbf{y}^* = W^{1/2}\mathbf{y}$, $X^* = W^{1/2}X$, and $\mathbf{e}^* = W^{1/2}\mathbf{e}$. One important feature after the transformation is that $\mathbf{e}^*$ is now distributed with a known identity matrix I_k :

$$\text{var}(\mathbf{e}^*) = W^{1/2}\,\text{var}(\mathbf{e})W^{1/2}$$

$$= W^{1/2}V_e W^{1/2} = I_k,$$

where $W = V_e^{-1}$.

Since the transformed error $\mathbf{e}^*$ is assumed to be independent and identically distributed (iid) with a unit variance, ordinary least squares (OLS) and ML method can be directly applied to the meta-analytic data. In other words, SEM may also be used to fit models on the transformed effect sizes.

Figure 1 shows a graphical model on the fixed-effects meta-analysis. Using conventional SEM notation, squares, circles, and triangles represent the observed variables, the latent variables, and the means, respectively. There are two points that require special attention. First, instead of estimating the error variance on $\mathbf{y}^*$, it is fixed as 1. Second, the intercept of $\mathbf{y}^*$ is fixed as 0 because the estimated population effect size is now represented by $\hat{\beta}_F$ (b0 in the figure). These two constraints are crucial in applying the SEM based meta-analysis.

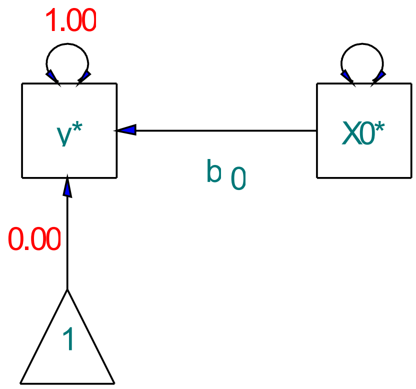

Figure 1

Random-effects models. A random-effects meta-analysis can be formulated as a single-level analysis with random slopes in SEM:

$$\mathbf{y}^* = I_k^* \mathbf{u} + \mathbf{e}^*,$$

where $\mathbf{u} \sim N(\beta_0 \mathbf{1}, I_k \tau^2)$. Figure 2 shows the graphical model in which $X_0^* = I_k^*$ is the transformed vector of ones. u from X_0^* to $\mathbf{y}^*$ represents a random slope that varies across studies. Thus, u_i in the ith study is treated as a random variable. It can be easily shown that the mean of u_i is the estimated mean effect size $\hat{\beta}_R$, while the variance of u_i (m in Figure 2) is the estimated variance component $\hat{\tau}^2$.

Since u_i varies across subjects (studies in the context of a meta-analysis), it is necessary to conduct a random slope analysis (Mehta & Neale, 2005; Muthén & Asparouhov, 2002, 2003; Muthén & Muthén, 2007).

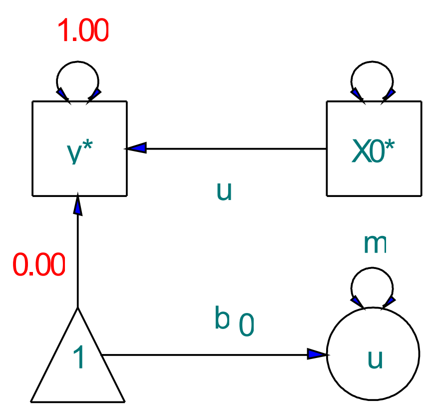

Figure 2

Mixed-effects models. The above transformation may also be applied to mixed-effects models. The mixed-effects model based on the transformed data is

$$W^{1/2}\mathbf{y} = W^{1/2}X\boldsymbol{\beta} + W^{1/2}I_k\mathbf{u} + W^{1/2}\mathbf{e}$$

$$\mathbf{y}^* = X^*\boldsymbol{\beta} + I_k^*\mathbf{u} + \mathbf{e}^*,$$

where $I_k^* = W^{1/2} I_k$. After the transformation, $\mathbf{e}^*$ is assumed to be distributed with a known identity matrix I_k. It should be noted that the same transformation with $W^{1/2}$ is applied regardless of whether the model is a fixed-, random- or mixed-effects one because the conditional variance σ_i^2 is the same under all models. Figure 3 shows the graphical model on a meta-regression where X1* is the transformed predictor.

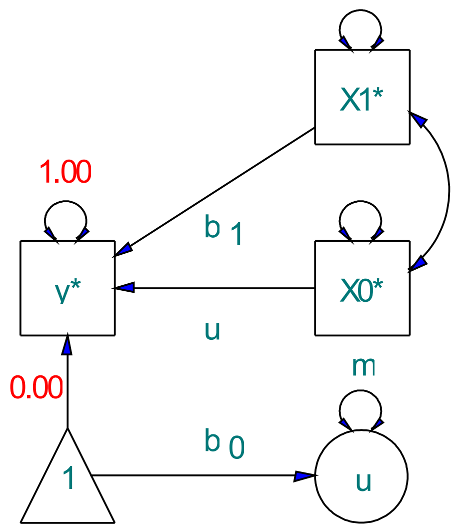

Figure 3

A data set from Hox (2002) is used to illustrate the procedures. Complete Mplus and Mx syntax and output are available here.

Data file (hox.txt): y (standardized mean difference) varofy (known variance of the effect size) inter (intercept) weeks (covariate)

```
-0.264 0.086 1 3
-0.230 0.106 1 1
 0.166 0.055 1 2
 0.173 0.084 1 4
 0.225 0.071 1 3
 0.291 0.078 1 6
 0.309 0.051 1 7
 0.435 0.093 1 9
 0.476 0.149 1 3
 0.617 0.095 1 6
 0.651 0.110 1 6
 0.718 0.054 1 7
 0.740 0.081 1 9
 0.745 0.084 1 5
 0.758 0.087 1 6
 0.922 0.103 1 5
 0.938 0.113 1 5
 0.962 0.083 1 7
 1.522 0.100 1 9
 1.844 0.141 1 9
```

```
TITLE: Fixed-effects model
```

```
DATA: FILE IS hox.txt;
VARIABLE: NAMES y varofy inter weeks;
      USEVARIABLES ARE y inter;     ! Inter: intercept of the model
DEFINE: w2 = SQRT(varofy**(-1));    ! Weight for the transformation
      y = w2*y;                     ! Transformed y
      inter = w2*inter;             ! Transformed intercept
MODEL:
      y ON inter;
      [y@0.0];                      ! Intercept is fixed at 0
      y@1.0;                        ! Error variance is fixed at 1
OUTPUT: SAMPSTAT;
```

Selected output:

```
MODEL RESULTS
                                                       Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
 Y        ON
    INTER              0.550      0.065      8.465      0.000
 Intercepts
    Y                  0.000      0.000    999.000    999.000
 Residual Variances
    Y                  1.000      0.000    999.000    999.000
```

Results:

1. The estimated population effect size under the fixed-effects model is 0.550 (*SE*=0.065) which is statistically significant at .05.

```
TITLE: Random-effects model
DATA: FILE IS hox.txt;
VARIABLE: NAMES y varofy inter weeks;
USEVARIABLES ARE y inter;     ! Inter: intercept of the model
DEFINE: w2 = SQRT(varofy**(-1));
      y = w2*y;
      inter = w2*inter;
ANALYSIS: TYPE=RANDOM;        ! Use random slope analysis
MODEL:
      [y@0.0];                ! Intercept is fixed at 0
      y@1.0;                  ! Error variance is fixed at 1
      u | y ON inter;         ! u: random effects
      u*;                     ! var(u): tau^2
      [u*];                   ! mean(u): estimated mean effect size
OUTPUT: SAMPSTAT;
```

Selected output:

```
MODEL RESULTS
                                                       Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value
 Means
    U                  0.579      0.107      5.406      0.000
 Intercepts
    Y                  0.000      0.000    999.000    999.000
 Variances
    U                  0.132      0.078      1.689      0.091
 Residual Variances
    Y                  1.000      0.000    999.000    999.000
```

Results:

1. The estimated mean effect size under the random-effects model is 0.579 (*SE*=0.107) which is statistically significant at .05.
2. The estimated variance component is 0.132.

```
TITLE: Meta-regression with a continuous covariate
DATA: FILE IS hox.txt;
VARIABLE: NAMES y varofy inter weeks;
USEVARIABLES ARE y inter weeks;
DEFINE: w2 = SQRT(varofy**(-1));
      y = w2*y;
      inter = w2*inter;
      weeks = w2*weeks;
ANALYSIS: TYPE=RANDOM;         ! Use random slope analysis
MODEL:
      [y@0.0];                 ! Intercept is fixed at 0
      y@1.0;                   ! Error variance is fixed at 1
      u | y ON inter;
      y ON weeks;
      u*;                      ! var(u): tau^2
      [u*];                    ! mean(u): intercept
OUTPUT: SAMPSTAT;
```

Selected output:

```
MODEL RESULTS
                                                    Two-Tailed
                    Estimate       S.E.  Est./S.E.    P-Value

 Y          ON
    WEEKS              0.139      0.036      3.904      0.000
 Means
    U                 -0.214      0.171     -1.246      0.213

 Intercepts
    Y                  0.000      0.000    999.000    999.000
 Variances
    U                  0.023      0.029      0.809      0.418
 Residual Variances
    Y                  1.000      0.000    999.000    999.000
```

Results:

1. The estimated regression coefficient from *weeks* to the effect size is 0.139 (*SE*=0.036) which is statistically significant at .05.
2. The estimated intercept is -0.214 (*SE*=0.171) which is not statistically significant at .05.
3. The estimated variance component is 0.023.

Besides the above illustrations, the SEM-based meta-analysis may also be used to handle missing covariates with ML method, to quantify the heterogeneity of effect sizes, and to address the heterogeneity of effect sizes with mixture models (see Cheung, 2008 for details).

Meta-analytic Structural Equation Modeling

Meta-analytic structural equation modeling (MASEM) refers to the techniques of synthesizing correlation (or covariance) matrices and fitting structural equation models to the pooled correlation (or covariance) matrix. Different researchers use different names for similar techniques, for example, meta-analytic path analysis, meta-analysis of factor analysis, path analysis of meta-analytically derived correlation matrices, structural equation modeling of a meta-analytic correlation matrix and path analysis based on meta-analytic findings.

Researchers typically conduct MASEM via a two-stage procedure (Cheung & Chan, 2005a, 2009; Viswesvaran & Ones, 1995). In the first stage, correlation matrices are tested for homogeneity. If they are not significantly different from each other, they are combined to form a pooled correlation matrix. In the second stage, the pooled correlation matrix is treated as the observed correlation matrix and used in SEM.

Methods to conducting MASEM

Univariate and multivariate approaches are available to conducting MASEM. The univariate methods are based on Hedge and Olkin (1985) or Hunter and Schmidt (1990) while the multivariate methods can be based on the generalized least squares (GLS; Becker, 1992, 1995; S. F. Cheung, 2000; Furlow &, Beretvas, 2005; Hafdahl, 2001) and SEM (Cheung & Chan, 2005a, 2009). Comparisons of some of these approaches can be found in Cheung and Chan (2005a).

Two-stage Structural Equation Modeling

Two-stage structural equation modeling (TSSEM) is an SEM approach proposed by Cheung (2002) and Cheung and Chan (2005a, 2009) to conduct MASEM. Simply put, multiple-group SEM is used to synthesize correlation matrices in the first stage. If the correlation matrices is homogeneous, the pooled correlation matrix is used as the input in fitting structural models. The asymptotic covariance matrix of the pooled correlation matrix is used as the weight matrix with the weighted least squares (WLS) method as the estimation method. The main difference between the TSSEM and the other approaches is that all Stage 1 and 2 analyses of the TSSEM approach are conducted under the general SEM framework.

If the correlation matrices are heterogeneous, it may not be appropriate to combine them. Cheung and Chan (2005b) suggested using cluster analysis to classify the studies into relatively homogeneous subgroups. The TSSEM approach has been illustrated with the following data sets-International Social Survey Program (1989) in Cheung and Chan (2005a, 2009), Digman (1997) in Cheung and Chan (2005b) and Social Axioms in Cheung, Leung, and Au (2006).

An illustration

To illustrate how to apply the TSSEM approach, a real data set on work-related attitudes was considered (International Social Survey Program, 1989). Persons aged 18 years and older from 11 countries were sampled based on multistage stratified probability sampling. Four countries were selected for illustration here (see Cheung & Chan, 2005a for the complete analysis on the 11 countries). Nine variables were selected for demonstration purposes. They were grouped into three meaningful constructs: Job prospects (F_1), including job security (x_1), income (x_2), and advancement opportunity (x_3); Job nature (F_2), including interesting job (x_4), independent work (x_5), help other people (x_6), and useful to society (x_7); and Time demand (F_3), represented by flexible working hours (x_8) and lots of leisure time (x_9).

The data file (ISSP.dat) is listed below. As a demonstration on how ML method may be used to handle missing data, some correlations were randomly deleted. The missing variances and the missing correlations are represented by 1 and 0 in the data file, respectively. A DOS program (Cheung, 2009) was used to generate the LISREL syntax and to handle the data manipulations between the Stage 1 and Stage 2 analyses.

```
1.00000
```

```
.32109 1.00000
.25886 .44019 1.00000
.30143 .30423 .31103 1.00000
.25063 .31368 .26286 .55888 1.00000
.21818 .18534 .22199 .43890 .37990 1.00000
.21270 .13287 .17488 .45450 .33126 .63329 1.00000
.05951 .17767 .14354 .07526 .26553 .11531 .10124 1.00000
.11967 .10506 .18103 .11833 .07252 .11642 .13303 -.01537 1.00000
1.00000
.00000 1.00000
.00000 .00000 1.00000
.00000 .00000 .30627 1.00000
.00000 .00000 .23286 .37229 1.00000
.00000 .00000 .16685 .40657 .24433 1.00000
.00000 .00000 .11537 .38946 .17179 .57266 1.00000
.00000 .00000 .20144 .15372 .21161 .14601 .13517 1.00000
.00000 .00000 .06605 .01053 .04781 .01423 .04271 .23570 1.00000
1.00000
.33292 1.00000
.00000 .00000 1.00000
.00000 .00000 .00000 1.00000
.00000 .00000 .00000 .00000 1.00000
.15641 .07923 .00000 .00000 .00000 1.00000
.17289 .07578 .00000 .00000 .00000 .52587 1.00000
.13045 .14124 .00000 .00000 .00000 .21323 .17417 1.00000
.10470 .13151 .00000 .00000 .00000 .04169 .04298 .31630 1.00000
1.00000
.23844 1.00000
.13480 .30850 1.00000
.16443 .22380 .30570 1.00000
.20362 .16293 .09850 .42835 1.00000
.20403 .10355 .14059 .34044 .31760 1.00000
.20217 -.04808 .10247 .25174 .19930 .50159 1.00000
.00000 .00000 .00000 .00000 .00000 .00000 .00000 1.00000
.00000 .00000 .00000 .00000 .00000 .00000 .00000 .00000 1.00000
```

Stage 1: Testing the homogeneity of correlation matrices. In the first stage, the homogeneity of correlation matrix is tested with multiple-group SEM.

LISREL syntax:

```
TI TSSEM Stage 1: Group 1
DA NG=4 NI=9 NO=591
CM SY FI=issp.dat
MO NX=9 NK=9 LX=DI,FR TD=ZE PH=ST,FR
OU
TI TSSEM Stage 1: Group 2
DA NI=9 NO=656
CM SY FI=issp.dat
MO NX=9 NK=9 LX=DI,FR TD=ZE PH=ST,FR
OU
TI TSSEM Stage 1: Group 3
DA NI=9 NO=832
CM SY FI=issp.dat
MO NX=9 NK=9 LX=DI,FR TD=ZE PH=ST,FR
OU
TI TSSEM Stage 1: Group 4
DA NI=9 NO=823
```

```
CM SY FI=issp.dat
MO NX=9 NK=9 LX=DI,FR TD=ZE PH=ST,FR
! Constraints on testing the homogeneity of correlation matrices
EQ PH(1,3,4) PH(2,3,4)
EQ PH(1,3,5) PH(2,3,5)
EQ PH(1,3,6) PH(2,3,6)
EQ PH(1,3,7) PH(2,3,7)
EQ PH(1,3,8) PH(2,3,8)
EQ PH(1,3,9) PH(2,3,9)
EQ PH(1,4,5) PH(2,4,5)
EQ PH(1,4,6) PH(2,4,6)
EQ PH(1,4,7) PH(2,4,7)
EQ PH(1,4,8) PH(2,4,8)
EQ PH(1,4,9) PH(2,4,9)
EQ PH(1,5,6) PH(2,5,6)
EQ PH(1,5,7) PH(2,5,7)
EQ PH(1,5,8) PH(2,5,8)
EQ PH(1,5,9) PH(2,5,9)
EQ PH(1,6,7) PH(2,6,7)
EQ PH(1,6,8) PH(2,6,8)
EQ PH(1,6,9) PH(2,6,9)
EQ PH(1,7,8) PH(2,7,8)
EQ PH(1,7,9) PH(2,7,9)
EQ PH(1,8,9) PH(2,8,9)
EQ PH(1,1,2) PH(3,1,2)
EQ PH(1,1,6) PH(3,1,6)
EQ PH(1,1,7) PH(3,1,7)
EQ PH(1,1,8) PH(3,1,8)
EQ PH(1,1,9) PH(3,1,9)
EQ PH(1,2,6) PH(3,2,6)
EQ PH(1,2,7) PH(3,2,7)
EQ PH(1,2,8) PH(3,2,8)
EQ PH(1,2,9) PH(3,2,9)
EQ PH(1,6,7) PH(3,6,7)
EQ PH(1,6,8) PH(3,6,8)
EQ PH(1,6,9) PH(3,6,9)
EQ PH(1,7,8) PH(3,7,8)
EQ PH(1,7,9) PH(3,7,9)
EQ PH(1,8,9) PH(3,8,9)
EQ PH(1,1,2) PH(4,1,2)
EQ PH(1,1,3) PH(4,1,3)
EQ PH(1,1,4) PH(4,1,4)
EQ PH(1,1,5) PH(4,1,5)
EQ PH(1,1,6) PH(4,1,6)
EQ PH(1,1,7) PH(4,1,7)
EQ PH(1,2,3) PH(4,2,3)
EQ PH(1,2,4) PH(4,2,4)
EQ PH(1,2,5) PH(4,2,5)
EQ PH(1,2,6) PH(4,2,6)
EQ PH(1,2,7) PH(4,2,7)
EQ PH(1,3,4) PH(4,3,4)
EQ PH(1,3,5) PH(4,3,5)
EQ PH(1,3,6) PH(4,3,6)
EQ PH(1,3,7) PH(4,3,7)
EQ PH(1,4,5) PH(4,4,5)
EQ PH(1,4,6) PH(4,4,6)
EQ PH(1,4,7) PH(4,4,7)
```

```
EQ PH(1,5,6) PH(4,5,6)
EQ PH(1,5,7) PH(4,5,7)
EQ PH(1,6,7) PH(4,6,7)
OU PH=cor1.cor EC=cor1.ack
```

Selected output:

```
                    Global Goodness of Fit Statistics

                         Degrees of Freedom = 57
              Minimum Fit Function Chi-Square = 172.73 (P = 0.00)
       Normal Theory Weighted Least Squares Chi-Square = 175.41 (P = 0.00)
               Estimated Non-centrality Parameter (NCP) = 118.41
           90 Percent Confidence Interval for NCP = (82.29 ; 162.16)

                    Minimum Fit Function Value = 0.060
               Population Discrepancy Function Value (F0) = 0.041
             90 Percent Confidence Interval for F0 = (0.028 ; 0.056)
             Root Mean Square Error of Approximation (RMSEA) = 0.054
           90 Percent Confidence Interval for RMSEA = (0.045 ; 0.063)
              P-Value for Test of Close Fit (RMSEA < 0.05) = 0.25

                Expected Cross-Validation Index (ECVI) = 0.15
             90 Percent Confidence Interval for ECVI = (0.13 ; 0.16)
                      ECVI for Saturated Model = 0.031
                     ECVI for Independence Model = 1.42

     Chi-Square for Independence Model with 144 Degrees of Freedom = 4094.05
                        Independence AIC = 4166.05
                            Model AIC = 421.41
                          Saturated AIC = 360.00
                       Independence CAIC = 4417.09
                           Model CAIC = 1279.11
                         Saturated CAIC = 1615.17

                      Normed Fit Index (NFI) = 0.96
                    Non-Normed Fit Index (NNFI) = 0.93
                 Parsimony Normed Fit Index (PNFI) = 0.38
                    Comparative Fit Index (CFI) = 0.97
                    Incremental Fit Index (IFI) = 0.97
                     Relative Fit Index (RFI) = 0.89

         PHI
               VAR 1      VAR 2      VAR 3      VAR 4      VAR 5      VAR 6
            --------   --------   --------   --------   --------   --------
     VAR 1      1.00

     VAR 2      0.30       1.00
              (0.02)
               15.34

     VAR 3      0.20       0.37       1.00
              (0.02)     (0.02)
                7.92      16.38

     VAR 4      0.22       0.25       0.31       1.00
              (0.02)     (0.02)     (0.02)
```

```
             9.06      10.79      15.51

   VAR 5     0.22       0.23       0.19       0.45       1.00
           (0.02)     (0.02)     (0.02)     (0.02)
             9.00       9.50       9.02      25.90

   VAR 6     0.19       0.11       0.17       0.39       0.31       1.00
           (0.02)     (0.02)     (0.02)     (0.02)     (0.02)
             9.31       5.58       8.06      20.93      15.98

   VAR 7     0.19       0.05       0.13       0.36       0.23       0.55
           (0.02)     (0.02)     (0.02)     (0.02)     (0.02)     (0.01)
             9.53       2.19       6.20      18.93      11.36      42.99

   VAR 8     0.10       0.15       0.17       0.12       0.24       0.16
           (0.03)     (0.03)     (0.03)     (0.03)     (0.03)     (0.02)
             3.84       5.88       6.23       4.50       9.21       7.65

   VAR 9     0.10       0.10       0.13       0.06       0.06       0.05
           (0.03)     (0.03)     (0.03)     (0.03)     (0.03)     (0.02)
             3.82       3.97       4.63       2.11       2.38       2.43

        PHI
            VAR 7      VAR 8      VAR 9
         --------   --------   --------
   VAR 7     1.00

   VAR 8     0.14       1.00
           (0.02)
             6.35

   VAR 9     0.07       0.20       1.00
           (0.02)     (0.02)
             3.05       9.39
```

Results:

1. The correlation matrices are quite homogeneous with $\chi^2(57) = 175.41, p < .001, RMSEA = 0.054, CFI = 0.97, NNFI = 0.93$.

Stage 2: Fitting a confirmatory factor analytic model. The pooled correlation matrix is then used to fit a factor analytic model.

LISREL syntax:

```
TI TSSEM Stage 2
DA NI=9 NO=2902 MA=KM
KM=cor2.cor SY
AC=cor2.ack SY
MO NX=9 NK=3 PH=ST,FR
PA LX
3*(1 0 0) 4*(0 1 0) 2*(0 0 1)
PD
OU ME=WL
```

Selected output:

```
 LISREL Estimates (Weighted Least Squares)
        LAMBDA-X
```

	KSI 1	KSI 2	KSI 3
VAR 1	0.52 (0.02) 21.32	- -	- -
VAR 2	0.57 (0.02) 23.75	- -	- -
VAR 3	0.59 (0.03) 22.16	- -	- -
VAR 4	- -	0.71 (0.02) 45.28	- -
VAR 5	- -	0.58 (0.02) 32.77	- -
VAR 6	- -	0.75 (0.01) 53.33	- -
VAR 7	- -	0.70 (0.01) 47.42	- -
VAR 8	- -	- -	0.61 (0.05) 11.56
VAR 9	- -	- -	0.33 (0.03) 9.97

PHI

	KSI 1	KSI 2	KSI 3
KSI 1	1.00		
KSI 2	0.54 (0.03) 21.13	1.00	
KSI 3	0.48 (0.05) 8.95	0.39 (0.04) 9.19	1.00

THETA-DELTA

VAR 1	VAR 2	VAR 3	VAR 4	VAR 5	VAR 6
0.73 (0.03) 23.55	0.68 (0.03) 20.79	0.65 (0.04) 17.92	0.50 (0.03) 17.44	0.66 (0.03) 23.88	0.44 (0.03) 15.82

```
THETA-DELTA
      VAR 7      VAR 8      VAR 9
   --------   --------   --------
       0.52       0.62       0.89
     (0.03)     (0.07)     (0.03)
      18.72       9.21      30.78

                    Goodness of Fit Statistics

                     Degrees of Freedom = 24
         Minimum Fit Function Chi-Square = 378.58 (P = 0.0)
         Estimated Non-centrality Parameter (NCP) = 354.58
     90 Percent Confidence Interval for NCP = (295.27 ; 421.33)

                 Minimum Fit Function Value = 0.13
         Population Discrepancy Function Value (F0) = 0.12
       90 Percent Confidence Interval for F0 = (0.10 ; 0.15)
      Root Mean Square Error of Approximation (RMSEA) = 0.071
     90 Percent Confidence Interval for RMSEA = (0.065 ; 0.078)
        P-Value for Test of Close Fit (RMSEA < 0.05) = 0.00

           Expected Cross-Validation Index (ECVI) = 0.14
      90 Percent Confidence Interval for ECVI = (0.12 ; 0.17)
                  ECVI for Saturated Model = 0.031
                 ECVI for Independence Model = 1.15

Chi-Square for Independence Model with 36 Degrees of Freedom = 3309.20
                    Independence AIC = 3327.20
                        Model AIC = 420.58
                      Saturated AIC = 90.00
                   Independence CAIC = 3389.96
                       Model CAIC = 567.02
                     Saturated CAIC = 403.79

                   Normed Fit Index (NFI) = 0.89
                 Non-Normed Fit Index (NNFI) = 0.84
              Parsimony Normed Fit Index (PNFI) = 0.59
                 Comparative Fit Index (CFI) = 0.89
                 Incremental Fit Index (IFI) = 0.89
                  Relative Fit Index (RFI) = 0.83
```

Results:

1. The proposed model fits the data marginally with $\chi^2(24) = 378.58, p < .001, RMSEA = 0.071, CFI = 0.89, NNFI = 0.84$. The parameter estimates are shown in Figure 4.

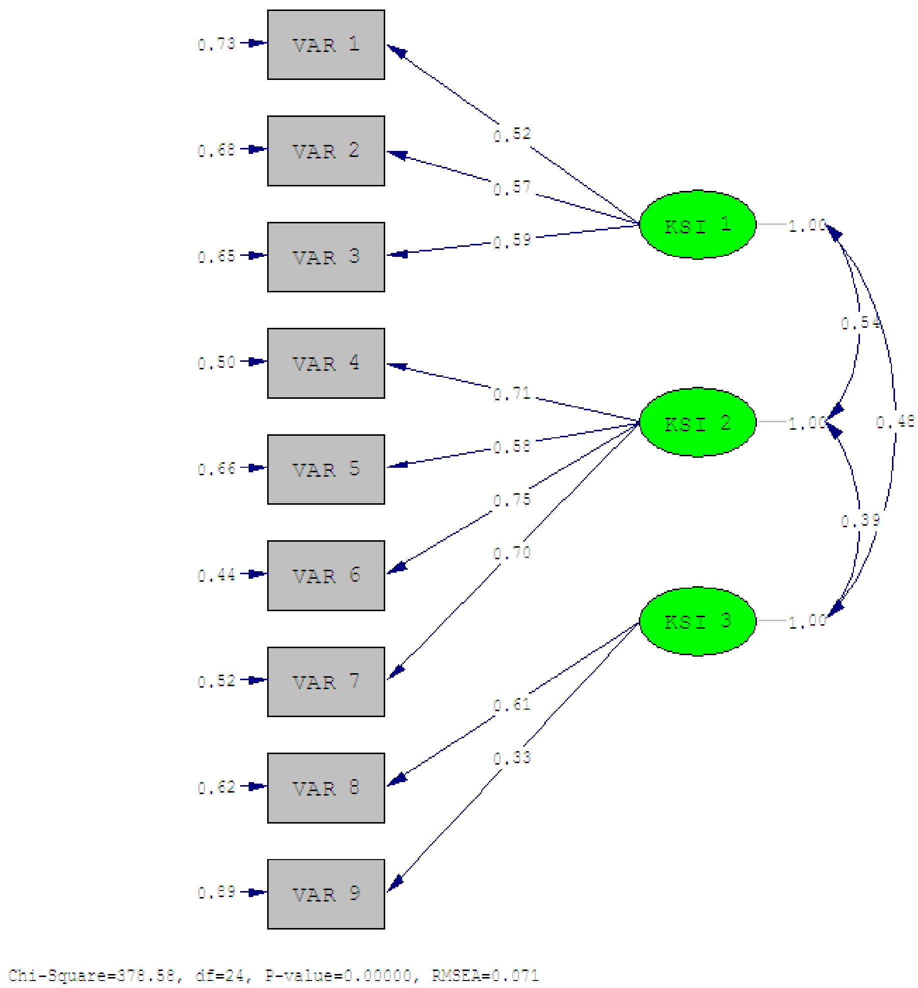

Figure 4

This note has demonstrated how SEM can be used to conduct fixed-, random-, and mixed-effects meta-analysis and MASEM. To summarize, SEM provides a flexible framework to conduct meta-analysis. As many state-of-the-art techniques have been implemented in the current SEM packages, it is expected that some of these techniques will prove useful to researchers conducting meta-analysis.

References

Becker, B. J. (1992). Using results from replicated studies to estimate linear models. *Journal of Educational Statistics, 17*, 341-362.

Becker, B. J. (1995). Corrections to "Using results from replicated studies to estimate linear models". *Journal of Educational Statistics*, *20*, 100-102.

Berkey, C. S., Hoaglin, D. C., Mostellar, F., & Colditz, G. A. (1995). A random effects regression model for meta-analysis. *Statistics in Medicine*, *14*, 395-411.

Cheung, M. W. L. (2002). *Meta-analysis for structural equation modeling: A two-stage approach.* Unpublished doctoral dissertation, Chinese University of Hong Kong, Hong Kong.

Cheung, M. W. L. (2008). A model for integrating fixed-, random-, and mixed-effects meta-analyses into structural equation modeling. *Psychological Methods*, *13*, 182-202. (PDF file and Mplus and Mx code)

Cheung, M. W. L. (2009). TSSEM: A LISREL syntax generator for two-stage structural equation modeling (Version 1.11) [Computer software]. Retrieved from http://courses.nus.edu.sg/course/psycwlm/internet/tssem.zip.

Cheung, M. W. L. (in press). Fixed-effects meta-analyses as multiple-group structural equation models. *Structural Equation Modeling*. (PDF file)

Cheung, M. W. L., & Chan, W. (2009). A two-stage approach to synthesizing covariance matrices in meta-analytic structural equation modeling. *Structural Equation Modeling*, *16*, 28-53.

Cheung, M. W. L., & Chan, W. (2005a). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, *10*, 40-64.

Cheung, M. W. L., & Chan, W. (2005b). Classifying correlation matrices into relatively homogeneous subgroups: A cluster analytic approach. *Educational and Psychological Measurement*, *65*, 954-979.

Cheung, M. W. L., Leung, K., & Au, K. (2006). Evaluating multilevel models in cross-cultural research: An Illustration with Social Axioms. *Journal of Cross-Cultural Psychology*, *37*, 522-541.

Cheung, S. F. (2000). *Examining solutions to two practical issues in meta-analysis: Dependent correlations and missing data in correlation matrices*. Unpublished doctoral dissertation, Chinese University of Hong Kong, Hong Kong.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 177-188.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246-1256.

Furlow, C. F. &, Beretvas, S. N. (2005). Meta-analytic methods of pooling correlation matrices for structural equation modeling under different patterns of missing data. *Psychological Method*s, *10*, 227-254.

Hafdahl, A. R. (2001). *Multivariate meta-analysis for exploratory factor analytic research*. Unpublished doctoral dissertation, University of North Carolina at Chapel Hill.

Hedges, L. V., & Vevea, J. L. (1998). Fixed- and random-effects models in meta-analysis. *Psychological Methods, 3*, 486-504.

Hox, J. J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage.

Hunter, J. E., & Schmidt, F. L. (2000). Fixed effects vs. random effects meta-analysis models: Implications for cumulative research knowledge. *International Journal of Selection and Assessment*, *8*, 275-292.

Inter-University Consortium for Political and Social Research (1989). *International Social Survey Program: Work Orientation*. Ann Arbor, MI: Author.

Kalaian, H. A., & Raudenbush, S. W. (1996). A multivariate mixed linear model for meta-analysis. *Psychological Methods*, *1*, 227-235.

Konstantopoulos, S. (2008). An introduction to meta-analysis. In J. W. Osborne (Ed.). *Best practices in quantitative methods* (pp. 177-194). Thousand Oaks, CA: Sage Publications.

Konstantopoulos, S., & Hedges, L. V. (2004). Meta-analysis. In M. Kaplan (Ed), *Handbook of quantitative methods in the social sciences* (pp. 281-297). Thousand Oaks, CA: Sage Publications.

Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equation modeling. *Psychological Methods*, *10*, 259-284.

Muthén, B, & Asparouhov, T. (2002). Modeling of heteroscedastic measurement errors. Retrieved from http://www.statmodel.com.

Muthén, B, & Asparouhov, T. (2003). Modeling interactions between latent and observed continuous variables using Maximum-Likelihood estimation In Mplus. Retrieved from http://www.statmodel.com.

Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide (5th ed.)*. Los Angeles, CA: Muthén, & Muthén.

National Research Council (1992). *Combining information: Statistical issues and opportunities for research*. Washington, D.C.: National Academy Press.

Raudenbush, S. W. (1994). Random effects models. In H. Cooper and L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 301-321). New York: Russell Sage Foundation.

Raudenbush, S. W., Becker, B. J., & Kalaian, H. (1988). Modeling multivariate effect sizes. *Psychological Bulletin*, *103*, 111-120.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd Ed.)*. Thousand Oaks, CA: Sage Publications.

Thompson, S. G., & Higgins, J. P. T. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*, 1559-1573.

Viechtbauer, W. (2005). Bias and efficiency of meta-analytic variance estimators in the random-effects model. *Journal of Educational and Behavioral Statistics*, *30*, 261-293.

Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, *48*, 865-885.