Wishart hierarchical models for meta-analytic latent variable models: A demonstration with the Hospital Anxiety and Depression Scale

James Ohisei Uanhoro

Research, Measurement & Statistics, Department of Educational Psychology, University of North Texas

Meta-analytic structural equation modeling

Meta-analytic structural equation modeling [MASEM, 2, 8] combines ideas from meta-analysis and SEM to estimate and test covariance structures assumed to underlie multiple covariance matrices. MASEMs are usually estimated under two assumptions:

1. Fixed-effects model: All covariance matrices have an identical population covariance matrix with observed differences due to sampling error.
2. Random-effects model: Covariance matrices have different population covariance matrices with observed differences due to both differences in population and sampling error.

Alternatively, multiple covariance matrices may be nested within a single study or author such that the covariance matrices are not independent of each other. There is **no systematic MASEM approach for handling dependent covariance matrices**.

Wishart-based MASEM solution

Under the assumption that the data in a study are multivariate normal, the $p \times p$ sample covariance matrix ($\mathbf{S}$) is Wishart:

$$n^*\mathbf{S} \sim \mathcal{W}_p(\boldsymbol{\Sigma}, n^*) \tag{1}$$

where n^* = sample size -1, $\boldsymbol{\Sigma}$ (scale matrix) is the population covariance matrix underlying the study. Assume $\boldsymbol{\Sigma}$ to be a structured covariance matrix, $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, e.g. $\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \boldsymbol{\Lambda\Phi\Lambda}' + \boldsymbol{\Theta}$ for a confirmatory factor model. Accordingly, the **fixed-effects MASEM** for k studies is:

$$n_i^*\mathbf{S}_i \sim \mathcal{W}_p(\boldsymbol{\Sigma}(\boldsymbol{\theta}),\ n_i^*) \ \text{ for } i \in \{1, \ldots, k\} \tag{2}$$

As an extension to equation 1, $\boldsymbol{\Sigma}$ may be assumed inverse-Wishart [9]:

$$\boldsymbol{\Sigma} \sim \mathcal{W}_p^{-1}(\boldsymbol{\Omega} \times m, m), \tag{3}$$

where $\boldsymbol{\Omega}$ is the true covariance matrix, and $m > p - 1$. Equations 1 and 3 form a hierarchical model for $\mathbf{S}$, resulting in a generalized matrix variate beta type II marginal distribution for $\mathbf{S}$ [9, 5]:

$$\mathbf{S} \sim \mathrm{GB}_p^{\mathrm{II}}\left(\frac{n^*}{2}, \frac{m}{2}, \frac{m}{n^*}\boldsymbol{\Omega}, \mathbf{0}_{p\times p}\right),\ \ln\mathcal{L} = f(p, m+n^*) - f(p,m) - f(p,n^*) + \frac{1}{2}\Bigg((n^* - p - 1)\ln|\mathbf{S}| + m\ln|\boldsymbol{\Omega}| - (n^* + m)\ln\left|\frac{m\boldsymbol{\Omega} + n^*\mathbf{S}}{m + n_i^*}\right|\Bigg) \tag{4}$$

where $f(p,x) = \ln\Gamma_p(x/2) - 0.5\,[xp\ln(x/2) - xp]$, and Γ_p is the multivariate gamma function [5, definition 1.4.2]. Assuming $\boldsymbol{\Omega}$ to be a structured covariance matrix, $\boldsymbol{\Omega}(\boldsymbol{\theta})$, leads to a **random-effects MASEM**:

$$\mathbf{S}_i \sim \mathrm{GB}_p^{\mathrm{II}}\left(\frac{n_i^*}{2}, \frac{m}{2}, \frac{m}{n_i^*}\boldsymbol{\Omega}(\boldsymbol{\theta}), \mathbf{0}_{p\times p}\right) \ \text{ for } i \in \{1, \ldots, k\} \tag{5}$$

Assuming j in $1, \ldots, c$ clusters of covariance matrices, the **dependent-samples MASEM** is:

$$\mathbf{S}_{ij} \sim \mathrm{GB}_p^{\mathrm{II}}\left(\frac{n_i^*}{2}, \frac{m_1}{2}, \frac{m_1}{n_i^*}\boldsymbol{\Psi}_{j[i]}, \mathbf{0}_{p\times p}\right) \ \text{ for } i \in \{1, \ldots, k\} \\ m_2\boldsymbol{\Psi}_j \sim \mathcal{W}(\boldsymbol{\Omega}(\boldsymbol{\theta}), m_2) \ \text{ for } j \in \{1, \ldots, c\} \tag{6}$$

where $\boldsymbol{\Psi}_j$ is an unstructured covariance matrix that varies by cluster j.

Notes about Wishart models

$\boldsymbol{\Sigma}(\boldsymbol{\theta})$ in equation 2, $\boldsymbol{\Omega}(\boldsymbol{\theta})$ in equation 5 and $\boldsymbol{\Omega}(\boldsymbol{\theta})$ in equation 6 are assumed to be the true covariance structure underlying the observed covariance matrices for their respective models.

For the random-effects model: $\varepsilon = (m - p + 1)^{-1/2}$, approximates the root mean square error of approximation (RMSEA) from assuming ($\boldsymbol{\Omega}(\boldsymbol{\theta})$) matches the different $\boldsymbol{\Sigma}_i$ [9].

For the dependent-samples model:

- $\varepsilon_{(1/2)} = (m_{(1/2)} - p + 1)^{-1/2}$, ε_1 and ε_2 are within- and between- RMSEA respectively.
- The total RMSEA, $\varepsilon = \left((m_1 - p + 1)^{-1} + (m_2 - p + 1)^{-1}\right)^{1/2}$.

Demonstration

The 14-item Hospital Anxiety and Depression scale [HADS, 10] is widely used to test distress in non-psychiatric patient populations. 28 correlation matrices of the HADS scale were collated and meta-analyzed by [6]. For demonstration, I compared **two theoretical configurations**:

1. Two **correlated factors**: anxiety (odd-numbered items) and depression (even-numbered items);
2. A **bifactor** model: general factor with anxiety and depression sub-factors; all uncorrelated.

I fit the **three Wishart methods** (fixed-effects, random-effects, dependent-samples) to both configurations above resulting in six estimated models. The 28 correlation matrices were clustered within 21 studies with the following cluster sizes: 1 (18 studies); 3 (2 studies); and 4 (1 study).

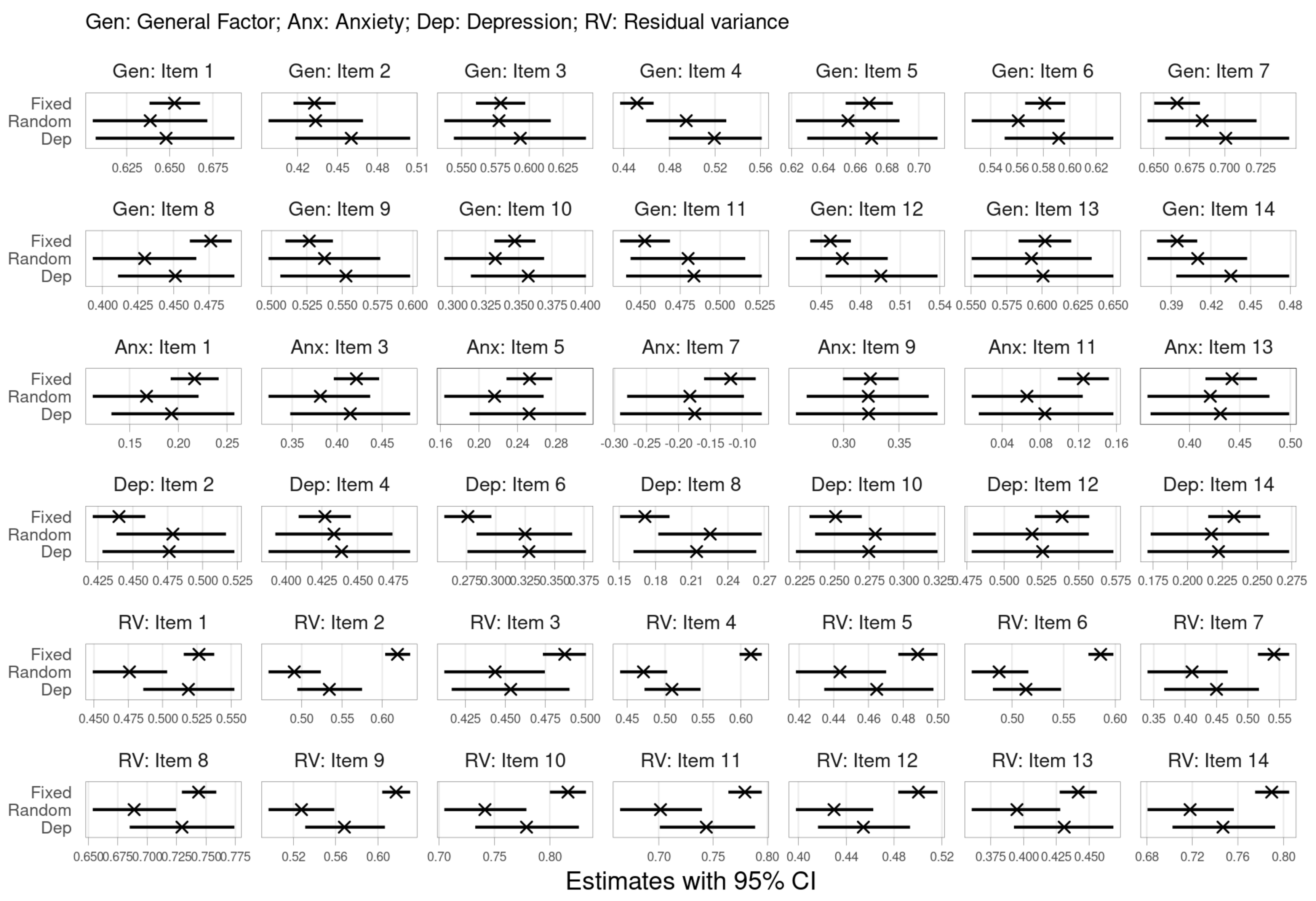
We applied Bayesian estimation using Stan [1], and LOOIC [7] for model comparison.

Table 1. Model comparison results sorted by LOOIC

Model	LOOIC	ΔLOOIC	Model weights
Dependent + bifactor	-8523.0	–	71.5%
Dependent + correlated	-8464.6	-58.4	25.4%
Random-effects + bifactor	-7025.2	-1439.4	3.0%
Random-effects + correlated	-6669.1	-356.1	$< 0.01\%$
Fixed-effects + bifactor	1699.6	-4969.5	$< 0.01\%$
Fixed-effects + correlated	4321.5	-2621.9	$< 0.01\%$

As with other commonplace information criteria, smaller values of LOOIC suggest better predictive performance of a model. The dependent-samples models had the best performance. Within any model type, the bifactor model was always the better model configuration. Accordingly, all additional results focus on the bifactor model.

Figure 1. Model estimates for bifactor models. Dependent-samples estimates have larger uncertainty. Fixed-effects model incorrectly assumes all covariance matrices have an identical population covariance matrix. Random-effects model incorrectly ignores non-independence of covariance matrices.



ε was 0.075, 95% CI [.072, .077] and 0.077, 95% CI [0.074, 0.079] for the random-effects and dependent-samples bifactor models respectively. Based on the dependent-samples model, much of the variance was between clusters ($\varepsilon_2 = 0.067$) as opposed to within clusters ($\varepsilon_1 = 0.040$). This accounts for the smaller uncertainty about estimates from the random-effects model in Figure 1.

References

[1] Bob Carpenter, Andrew Gelman, Matthew D Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Marcus A Brubaker, Peter Li, and Allen Riddell. Stan: A probabilistic programming language. *Journal of Statistical Software*, 76(1), 2017.

[2] Mike W.-L. Cheung and Wai Chan. Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, 10(1):40–64, March 2005.

[3] Mike W.-L. Cheung and Wai Chan. A two-stage approach to synthesizing covariance matrices in meta-analytic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(1):28–53, January 2009.

[4] Mike W.-L. Cheung and Ranjith Vijayakumar. A guide to conducting a meta-analysis. *Neuropsychology Review*, 26(2), 2016.

[5] A. K. Gupta and D. K. Nagar. *Matrix Variate Distributions*. CRC Press, October 1999.

[6] Sam Norton, Theodore Cosco, Frank Doyle, John Done, and Amanda Sacker. The hospital anxiety and depression scale: A meta confirmatory factor analysis. *Journal of Psychosomatic Research*, 74(1):74–81, January 2013.

[7] Aki Vehtari, Andrew Gelman, and Jonah Gabry. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432, 2017.

[8] Chockalingam Viswesvaran and Deniz S. Ones. Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, 48(4):865–885, 1995.

[9] Hao Wu and Michael W. Browne. Quantifying adventitious error in a covariance structure as a random effect. *Psychometrika*, 80(3):571–600, September 2015.

[10] Anthony S Zigmond and R Philip Snaith. The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67(6):361–370, 1983.

minorbsem R package: jamesuanhoro.github.io/minorbsem

Open Science Framework: https://osf.io/rstzk/

James.Uanhoro@unt.edu