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

Abstract

I present an approach to meta-analytic SEM that relies on hierarchical modeling of sample covariance matrices under the assumption that the matrices are Wishart. The approach handles the commonplace fixed- and random- effects meta-analytic SEMs, and solves the problem of dependent covariance matrices where more than one covariance matrix is obtained from a single study. The estimation approach is Bayesian, and I provide some guidance on prior specification, as well as model code to aid application and further study of the approach. Finally, I demonstrate the approach with 28 correlation matrices collected from 21 studies of the Hospital Anxiety and Depression scale.

Notation:

$\mathcal{N}()$: Normal distribution, mean-scale notation

$\mathcal{W}_p()$: Wishart distribution for a $p \times p$ matrix

$t^+()$: half-t distribution, degrees of freedom, location, scale notation

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

Structural equation modeling (SEM) is a popular statistical method for modeling covariance structures; for commonplace SEMs, the covariance matrix is sufficient for data analysis. Meta-analytic SEM (MASEM, Cheung & Chan, 2005; Viswesvaran & Ones, 1995) combines ideas from meta-analysis (Hedges & Olkin, 1985) and SEM to estimate and test covariance structures assumed to underlie multiple covariance matrices.

In this paper, I present an approach to MASEM based on the Wishart distribution. I specifically focus on latent-variable models, thus excluding meta-analytic path models. The approach includes both fixed- and random-effects MASEM, and additionally accounts for dependent covariance matrices. Prior to elaborating the approach, I briefly review the most popular approach to estimating MASEMs.

The most popular approach to estimating MASEMs is two-stage SEM (TSSEM, Cheung & Chan, 2005, 2009). In the first step, a pooled correlation matrix is estimated. Each $\mathbf{S}_i$ (for i in $1, \ldots, k$ sample covariance matrices) is decomposed thus: $\mathbf{R}_i = \mathbf{D}_i^{-1}\mathbf{S}_i\mathbf{D}_i^{-1}$, where $\mathbf{D}_i$ is a diagonal matrix with j^{th} diagonal element $s_{i(jj)}^{1/2}$, such that $\mathbf{R}_i$ is the correlation matrix. Under a fixed-effect TSSEM, all studies are assumed to have the same $\mathbf{R}_i$, i.e. $\mathbf{R}_1 = \ldots = \mathbf{R}_k$. Under a random-effects TSSEM, $\mathbf{R}_i$ is assumed to vary across studies. In a fixed-effect TSSEM, one can compute the degree of misspecification due to assuming the same $\mathbf{R}_i$ across studies. If the degree of misspecification is deemed unacceptable, then a random-effects TSSEM is preferred. Regardless of fixed-/random-effects approach, the resulting pooled correlations $\boldsymbol{\rho}$ are then used to estimate the hypothesized correlation structure $\boldsymbol{\rho}(\boldsymbol{\theta})$ with weighted least squares (WLS) discrepancy function (Browne, 1974). Other approaches to MASEM (e.g. one-step maximum-likelihood (Oort & Jak, 2016)) are often similar or asymptotically equivalent to TSSEM (Yuan & Kano, 2018). An exception is parameter-based MASEM (e.g. Ke, Zhang, & Tong, 2019) where structural parameters are assumed to vary across studies.

In the next section of the paper, I lay out the Wishart methods for MASEM. Afterward, I present a data analysis example using the Hospital Anxiety and Depression scale (HADS, Zigmond & Snaith, 1983) – the example shows how to practically apply the Wishart approach. I then conclude with discussion of the approach and some thoughts for further developing the approach. Finally, code for simulation studies and data analysis examples is available at `https://osf.io/rstzk/?view_only=61826849a2e94f978522a01d9b7a5a2e`.

Wishart-based MASEM

A major distinguishing feature of Wishart-based MASEM is that one meta-analyzes the covariance matrices directly when they are available, as opposed to the correlation matrices as typical in TSSEM. Under the assumption that the data in a study are multivariate normal, the sample covariance matrix is a Wishart variate:

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

where n^* = sample size $-$ 1, $\mathbf{\Sigma}$ (scale matrix) is the population covariance matrix underlying the study. As $n^* \to \infty,\ \mathbf{S} \to \mathbf{\Sigma}$, and the effect of sampling error is negligible.

Fixed-effects model

The sum of several Wishart matrices that share a common scale matrix is itself Wishart with the same scale matrix (Gupta & Nagar, 1999, Theorem 3.3.8):

$$\sum_{i=1}^{k}\left[n_i^*\mathbf{S}_i\right] \sim \mathcal{W}_p\left(\mathbf{\Sigma},\ \sum_{i=1}^{k} n_i^*\right) \text{ if } (\mathbf{S}_i, \ldots, \mathbf{S}_k) \sim \mathcal{W}_p\left(\frac{1}{n_i^*}\mathbf{\Sigma}, n^*\right), \tag{2}$$

where $\mathbf{\Sigma}$ in equation 2 is the pooled covariance matrix under a fixed-effects model. $\mathbf{\Sigma}$ may be further assumed to be a structured covariance matrix, $\mathbf{\Sigma}(\boldsymbol{\theta})$, such that one directly estimates the SEM parameters, $\boldsymbol{\theta}$. Hence, this would be a one-step fixed-effects MASEM. For meta-analytic confirmatory factor analysis (CFA) (the most common MASEM), the one-step fixed-effects MASEM would be:

$$\sum_{i=1}^{k}\left[n_i^*\mathbf{S}_i\right] \sim \mathcal{W}_p\left(\mathbf{\Lambda\Phi\Lambda}' + \mathbf{\Theta},\ \sum_{i=1}^{k} n_i^*\right), \tag{3}$$

where $\mathbf{\Lambda}$ is the loading matrix with certain elements set to 0 based on model identification and substantive considerations, $\mathbf{\Phi}$ is the inter-factor correlation matrix and $\mathbf{\Theta}$ is the residual covariance matrix.

Random-effects model

Given equation 1, the population covariance matrix, $\mathbf{\Sigma}$, may then be assumed to be inverse-Wishart (on the basis of conjugacy):

$$\mathbf{\Sigma} \sim \mathcal{W}_p^{-1}(\mathbf{\Omega} \times m, m), \tag{4}$$

where $\mathbf{\Omega}$ is the true covariance matrix underlying the population covariance matrix and $m > p - 1$ is the degrees of freedom and functions as a precision parameter – as $m \to \infty,\ \mathbf{\Sigma} \to \mathbf{\Omega}$. Wu and Browne (2015, hereafter WB) assumed $\mathbf{\Omega}$ to be a structured covariance matrix, $\mathbf{\Omega}(\boldsymbol{\theta})$, such that differences between $\mathbf{\Sigma}$ and $\mathbf{\Omega}(\boldsymbol{\theta})$ are due to what WB (2015) term *adventitious error* – error that arises because the exact study population differs from the hypothetical population for which the psychometric theory holds.

The models in equations 1 and 4 form a hierarchical model for $\mathbf{S}$ – the primary interest is in obtaining $\mathbf{\Omega}$ not $\mathbf{\Sigma}$. Hence, one can integrate out $\mathbf{\Sigma}$, the resulting marginal distribution for $\mathbf{S}$ is a generalized matrix variate beta type II (GMB-II) distribution (Roux & Becker, 1984, as cited in Wu & Browne, 2015):[1]

$$\mathbf{S} \sim \mathrm{GB}_p^{\mathrm{II}}\left(\frac{n^*}{2}, \frac{m}{2}, \frac{m}{n^*}\mathbf{\Omega}, \mathbf{0}_{p\times p}\right), \tag{5}$$

with log-likelihood:

[1] WB (2015) refer to this distribution as the second type of matrix variate beta distribution citing chapter 5 of Gupta and Nagar (1999). However, Gupta and Nagar (1999) in definition 5.2.4 include the label *generalized* to describe this distribution.

$$\ln \mathcal{L} = f(p, m + n^*) - f(p, m) - f(p, n^*) + \frac{1}{2}\Bigg(\\ (n^* - p - 1)\ln|\mathbf{S}| + m\ln|\mathbf{\Omega}| - (n^* + m)\ln\left|\frac{m\mathbf{\Omega} + n^*\mathbf{S}}{m + n_i^*}\right|\Bigg), \tag{6}$$

where $f(p, x) = \ln \Gamma_p(x/2) - \frac{1}{2}\left[xp\ln(x/2) - xp\right]$, and Γ_p is the multivariate gamma function (Gupta & Nagar, 1999, definition 1.4.2).

WB (2015) presented the model above in the context of a single study and suggested it may be useful for multi-group contexts. In MASEM contexts, it is reasonable to assume that differences in sample covariance matrices ($\mathbf{S}_i$) across studies result from different population covariance matrices ($\mathbf{\Sigma}_i$), with each of these populations being somewhat different from the generic population the hypothesized covariance structure holds for – if it holds.

The model in equation 5 is a hierarchical model for $\mathbf{S}$, I extend it to a random-effect MASEM for $\mathbf{S}_i$:

$$\mathbf{S}_i \sim \mathrm{GB}_p^{\mathrm{II}}\left(\frac{n_i^*}{2}, \frac{m}{2}, \frac{m}{n_i^*}\mathbf{\Omega}(\boldsymbol{\theta}), \mathbf{0}_{p\times p}\right) \text{ for } i \in \{1, \ldots, k\} \tag{7}$$

where $\mathbf{\Omega}(\boldsymbol{\theta})$ ($= \mathbf{\Lambda}\mathbf{\Phi}\mathbf{\Lambda}' + \mathbf{\Theta}$) is the pooled structured covariance matrix.

Based on results in WB (2015), the quantity, $v = 1/m$, represents the average variance between the study population covariance matrices. Additionally, the quantity, $1/\sqrt{(m + p - 1)}$, approximates the root mean square error of approximation (RMSEA, ε) from assuming the pooled structured ($\mathbf{\Omega}(\boldsymbol{\theta})$) matches the different $\mathbf{\Sigma}_i$ (Wu & Browne, 2015).

Dependent-samples model

Dependent covariance matrices can occur in MASEM when more than one covariance matrix is obtained from the same study. Commonplace MASEM approaches fail to account for such dependence. As a solution, I extend the model in equation 5. Assuming j in $1, \ldots, c$ clusters of covariance matrices, the following hierarchical model for dependent covariance matrices applies:

$$\begin{aligned} \mathbf{S}_{ij} &\sim \mathrm{GB}_p^{\mathrm{II}}\left(\frac{n_i^*}{2}, \frac{m_1}{2}, \frac{m_1}{n_i^*}\boldsymbol{\Psi}_{j[i]}, \mathbf{0}_{p\times p}\right) \text{ for } i \in \{1, \ldots, k\} \\ m_2\boldsymbol{\Psi}_j &\sim \mathcal{W}\left(\boldsymbol{\Omega}(\boldsymbol{\theta}), m_2\right) \text{ for } j \in \{1, \ldots, c\} \end{aligned} \tag{8}$$

where $\boldsymbol{\Psi}_j$ is an unstructured covariance matrix that varies by cluster j. $\boldsymbol{\Psi}_j$ is assumed Wishart (on the basis of conjugacy) with a scale parameter that is the true structured covariance matrix, $\boldsymbol{\Omega}(\boldsymbol{\theta})$. Hence, this remains a random-effect MASEM; the within-cluster variation is controlled by m_1 (precisely: $v_1 = 1/m_1$) and the between-cluster variation is controlled by m_2 ($v_2 = 1/m_2$); such that $v_1 + v_2 = v$, where v is the variation between population covariances and the pooled structured covariance matrix in the standard random-effects model, i.e. ignoring clustering. Additionally, the overall (ε), within-cluster (ε_1) and between-cluster (ε_2) RMSEAs can be computed, $\varepsilon_{(1/2)} = 1/\sqrt{m_{(1/2)} + p - 1}$. Finally, the proportion of variance that is between-cluster may also be of interest, $v_2(v_1 + v_2)^{-1}$.

Notes on the Wishart methods

The Wishart approach makes clear the data generation process for MASEM and sheds light on the nature of the pooled structured covariance matrix, $\boldsymbol{\Omega}(\boldsymbol{\theta})$. The approach assumes $\boldsymbol{\Omega}(\boldsymbol{\theta})$ underlies the observed covariance matrices. In the fixed-effects case, only sampling error is responsible for differences between the observed covariance matrices. In the random-effects case, adventitious error or study-specific context e.g. non-random

sampling of cases generates an intermediate population covariance matrix (true for the specific population sampled) between $\boldsymbol{\Omega}(\boldsymbol{\theta})$ and the observed covariance matrix. In the dependent-samples case, clustering creates another level of variation. Hence, random and non-random sampling variation may be responsible for $\boldsymbol{\Omega}(\boldsymbol{\theta})$ seeming very different from the observed $\mathbf{S}_i$.

Additionally, the Wishart methods as presented exploit conjugate pairings (Wishart $\rightarrow$ inverse-Wishart $\rightarrow$ Wishart) to account for higher levels of hierarchies (fixed $\rightarrow$ random $\rightarrow$ dependent).

Finally, the elaboration above focuses on the CFA, however, the methods are valid for any SEMs for which the covariance matrix is sufficient.

Implementation details

I assume a Bayesian model estimation approach and provide R and Stan (Carpenter et al., 2017) code to estimate the models above. Stan is a statistical programming language for fitting models using Markov Chain Monte Carlo methods using the relatively efficient No-U-Turn sampler.

Data analysis example

The data are 28 correlation matrices of the 14-item Hospital Anxiety and Depression scale (HADS, Zigmond & Snaith, 1983), collated and meta-analyzed by Norton, Cosco, Doyle, Done, and Sacker (2013) and re-analyzed by Jak and Cheung (2020). The HADS is widely used to test distress in non-psychiatric patient populations, and has been validated using factor analysis and item-response theory. For demonstration, I selected two competing theoretical configurations for the HADS:

1. Two correlated factors: anxiety (odd-numbered items) and depression (even-numbered items);

2. A bifactor model with anxiety and depression factors hypothesized as above and a general factor; all three factors are uncorrelated.

I fit each of the three Wishart methods (fixed-effects, random-effects, dependent-samples) to both theoretical 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).

I set the following priors for structural parameters across all three Wishart methods:

$$\boldsymbol{\lambda} \sim \mathcal{N}(0, \sigma_\lambda),\ \sigma_\lambda \sim t^+(3, 0, 1),\ \frac{\rho + 1}{2} \sim \text{Beta}(2, 2),\ \boldsymbol{\sigma}_\Theta \sim t^+(3, 0, 1) \tag{9}$$

where non-zero loadings ($\boldsymbol{\lambda}$) have a normal prior with scale hyperparameter, σ_λ. Both σ_λ and residual standard deviations ($\boldsymbol{\sigma}_\Theta$) have half-t priors (Gelman, 2006) creating a weakly regularizing effect (Lemoine, 2019) given that the total variable variances are 1. The prior for the interfactor correlation (ρ) is boundary-avoiding. Loadings and the interfactor correlation are sign-corrected post-sampling (e.g. Conti, Frühwirth-Schnatter, Heckman, & Piatek, 2014; Merkle, Fitzsimmons, Uanhoro, & Goodrich, 2021) to correct for rotational indeterminacy of latent variables (Peeters, 2012).

I assume the variance parameters in the random-effects method (v) and both within- (v_1) and between- variance parameters (v_2) in the dependent-samples method are standard half-normal, $[v, v_1, v_2] \sim \mathcal{N}^+(0, 1)$. This prior also has a weakly regularizing effect since the upper limit of these parameters is 0.077 ($1/(14 - 1)$) given the definition of the Wishart distribution.[2]

For each model, there were 1000 warmup iterations then 1000 iterations retained for inference across 4 chains. Sampler-specific diagnostics (Betancourt, 2017) were adequate for all estimated models. Across all models and parameters, the maximum $\widehat{R}$ (Vehtari, Gelman, Simpson, Carpenter, & Bürkner, 2020) was 1.02 suggesting parameter convergence for all parameters across models.

I compared the six estimated models using approximate leave-one-out information

[2] These variance parameters are inverse degrees of freedom for the Wishart distribution; the degrees of freedom parameter for a $p \times p$ covariance matrix must exceed $p - 1$.

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%

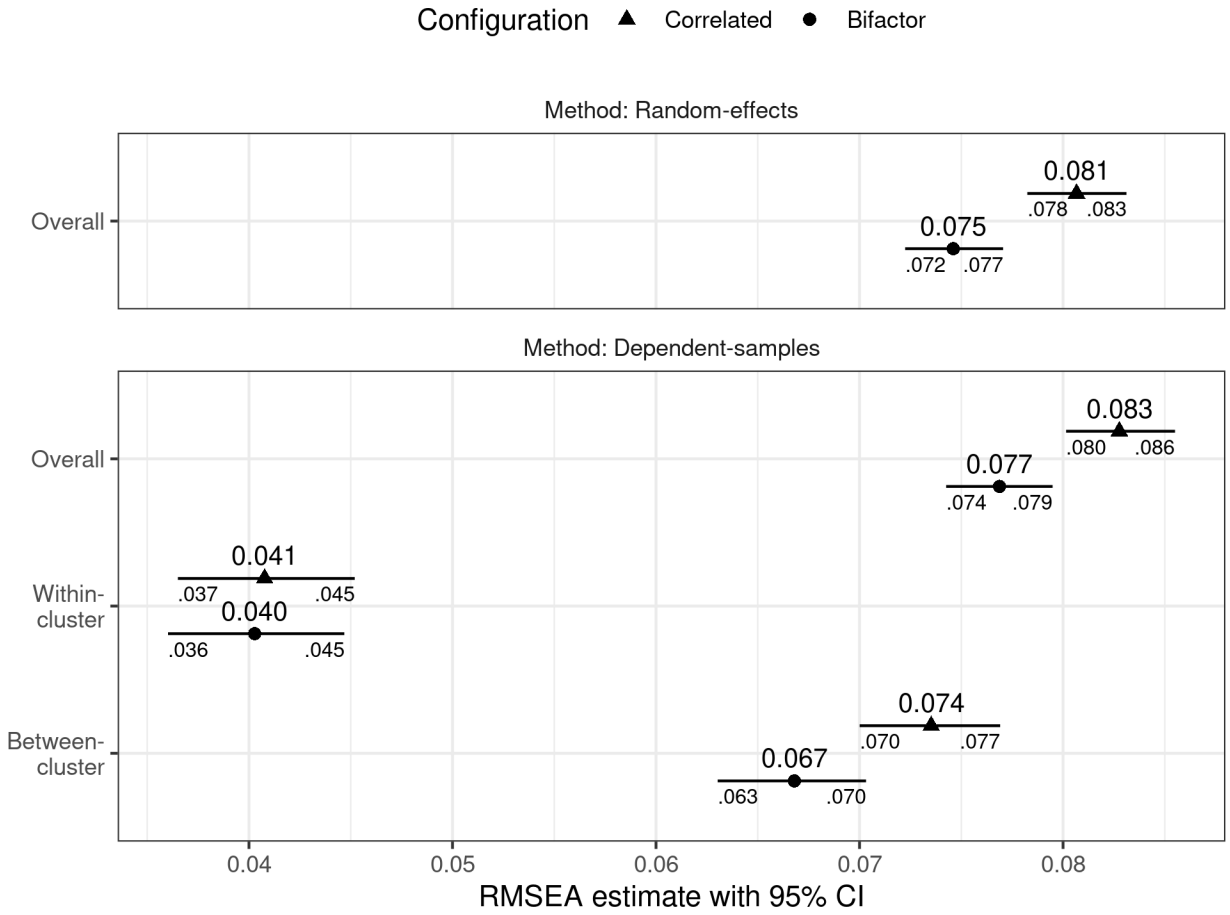
As with other commonplace information criteria, smaller values of LOOIC suggest better predictive performance of a model.

criterion (LOOIC, Vehtari, Gelman, & Gabry, 2017) and computed model weights via stacking (Yao, Vehtari, Simpson, & Gelman, 2018) using the *loo* package (Vehtari, Gabry, et al., 2020) in R. The worst models were the fixed-effects models, while the dependent-samples models were best performing, see Table 1. Given the same method, the bifactor configuration always outperformed the correlated factors configuration.

The RMSEA can also be used for comparing the random-effects and dependent-samples models, see Figure 1. For the random-effects models, the bifactor model had the lower RMSEA, suggesting that the average distance between the pooled structured covariance matrix and the population covariance matrices underlying individual observed covariance matrices was lower for the bifactor model compared to the correlated factors model.

There are three sets of RMSEA values for the dependent-samples models. The dependent-samples overall RMSEAs should match the random-effects RMSEAs – both sets of metrics capture the same information. The within-cluster RMSEAs (ε_1) should be identical across different model configurations of the same data is this is the gap between the cluster-level population covariance matrices and the population covariance matrices underlying each observed covariance matrix. Finally, the between-cluster RMSEA (ε_2) shows how the pooled structured covariance matrix differs from the cluster-level population covariance matrices, and this will differ by model configuration. Given that clustering of

Figure 1
Model RMSEA estimates

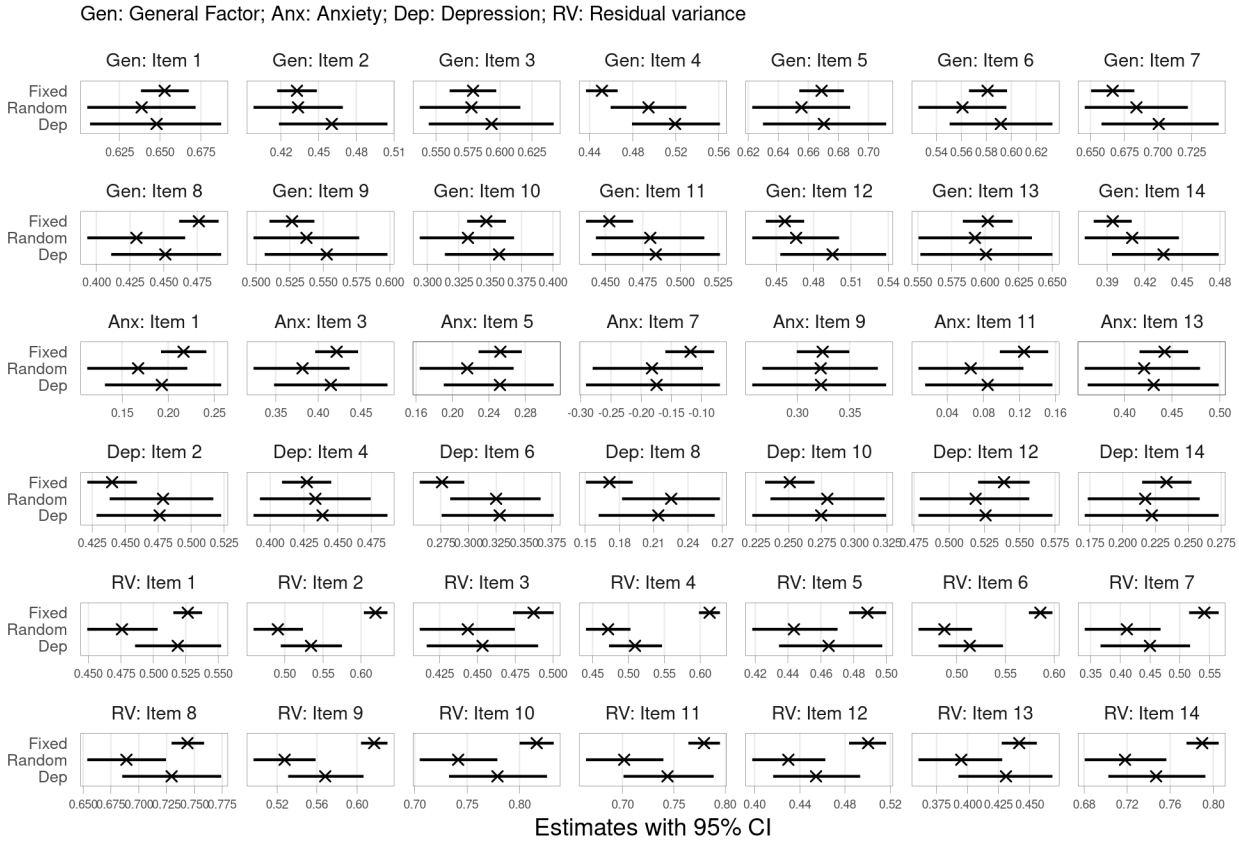


Note. Both panels have the same x-axis.

covariance matrices will create some shared variation, ε_2 is the best measure of how well the pooled structured covariance matrix differs from the individual population covariance matrices, as this metric reflects the RMSEA assuming there was no clustering. Based on this metric, the pooled bifactor model is less distant than the pooled correlated factors model from the individual population covariance matrices underlying observed covariance matrices. Hence, of the two configurations, the bifactor model better reflects the patterns responsible for generating the observed covariance matrices. That said, a considerable proportion of the variation is between clusters as opposed to within clusters for the bifactor model, 74.1%, 95% CI [68.0%, 79.4%]. This is unsurprising in this example since most

Figure 2

Estimates from multiple bifactor model solutions for HADS scale



clusters (18 / 21) are single-case clusters.

I report the parameter estimates of the three bifactor models in Figure 2. The pattern as one might predict is that accounting for more hierarchies – fixed-effects → random-effects and random-effects → dependent-samples – increases uncertainty about structural parameters. Alternatively stated, ignoring dependent-samples wrongly reduces the uncertainty about estimated structural parameters.

Discussion

In this paper, I have presented an approach to MASEM that uses hierarchical models based on the Wishart distribution to capture the two most common MASEM models (fixed- and random- effects) and solves an extant problem in the MASEM literature (dependent covariance matrices). The approach has been demonstrated with a dataset, the

demonstration showed how information-criteria based model comparison and the RMSEA produced by the models may be used for model comparison or selection.

A key limitation of the current work as presented is the absence of simulation results showing that the method returns adequate parameter estimates under a true model. I hope to address this in the future. In the meantime, that the method simply exploits conjugate pairings should increase confidence in the logical validity of the approach.

Another limitation of the proposed approach is the absence of global model fit indices that are commonplace within SEM. Specifically, the current method does not have a way of identifying non-trivial residual correlations induced by *minor* factors. To solve this problem, we propose estimating a full residual covariance matrix (similar to equation 9 in Muthén & Asparouhov, 2012) to allow for investigation of local misspecifications.

References

Betancourt, M. (2017, January). *A conceptual introduction to Hamiltonian Monte Carlo.* (arXiv: 1701.02434)

Browne, M. W. (1974, January). Generalized least squares estimators in the analysis of covariance structures. *South African Statistical Journal*, *8*(1), 1–24. doi: 10.10520/AJA0038271X_175

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1). doi: 10.18637/jss.v076.i01

Cheung, M. W.-L., & Chan, W. (2005, March). Meta-analytic structural equation modeling: A two-stage approach. *Psychological Methods*, *10*(1), 40–64. doi: 10.1037/1082-989X.10.1.40

Cheung, M. W.-L., & Chan, W. (2009, January). A two-stage approach to synthesizing covariance matrices in meta-analytic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(1), 28–53. doi: 10.1080/10705510802561295

Conti, G., Frühwirth-Schnatter, S., Heckman, J. J., & Piatek, R. (2014). Bayesian exploratory factor analysis. *Journal of Econometrics*, *183*(1), 31–57. doi: 10.1016/j.jeconom.2014.06.008

Gelman, A. (2006, September). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, *1*(3), 515–534. Retrieved from `http://projecteuclid.org/euclid.ba/1340371048` (Publisher: International Society for Bayesian Analysis) doi: 10.1214/06-BA117A

Gupta, A. K., & Nagar, D. K. (1999). *Matrix variate distributions.* CRC Press.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.

Jak, S., & Cheung, M. W. L. (2020). Meta-analytic structural equation modeling with

moderating effects on SEM parameters. *Psychological Methods*, *25*(4), 430–455. (Place: US Publisher: American Psychological Association) doi: 10.1037/met0000245

Ke, Z., Zhang, Q., & Tong, X. (2019, May). Bayesian meta-analytic SEM: A one-stage approach to modeling between-studies heterogeneity in structural parameters. *Structural Equation Modeling: A Multidisciplinary Journal*, *26*(3), 348–370. doi: 10.1080/10705511.2018.1530059

Lemoine, N. P. (2019). Moving beyond noninformative priors: Why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, *128*(7), 912–928. doi: 10.1111/oik.05985

Merkle, E. C., Fitzsimmons, E., Uanhoro, J., & Goodrich, B. (2021, November). Efficient Bayesian structural equation modeling in Stan. *Journal of Statistical Software*, *100*(6), 1–22. doi: 10.18637/jss.v100.i06

Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: A more flexible representation of substantive theory. *Psychological Methods*, *17*(3), 313–335. doi: 10.1037/a0026802

Norton, S., Cosco, T., Doyle, F., Done, J., & Sacker, A. (2013, January). The hospital anxiety and depression scale: A meta confirmatory factor analysis. *Journal of Psychosomatic Research*, *74*(1), 74–81. doi: 10.1016/j.jpsychores.2012.10.010

Oort, F. J., & Jak, S. (2016). Maximum likelihood estimation in meta-analytic structural equation modeling. *Research Synthesis Methods*, *7*(2), 156–167. doi: 10.1002/jrsm.1203

Peeters, C. F. W. (2012). Rotational uniqueness conditions under oblique factor correlation metric. *Psychometrika*, *77*(2), 288–292. Retrieved from `https://doi.org/10.1007/s11336-012-9259-3` doi: 10.1007/s11336-012-9259-3

Roux, J. J. J., & Becker, P. J. (1984). *On prior inverted Wishart distribution.* (Tech. Rep.). Pretoria, South Africa: Department of Statistics and Operations Research,

University of South Africa.

Vehtari, A., Gabry, J., Magnusson, M., Yao, Y., Bürkner, P.-C., Paananen, T., & Gelman, A. (2020). *loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models.* Retrieved from `https://mc-stan.org/loo/`

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432. Retrieved from `https://doi.org/10.1007/s11222-016-9696-4` doi: 10.1007/s11222-016-9696-4

Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. (2020). Rank-normalization, folding, and localization: An improved R for assessing convergence of MCMC. *Bayesian Analysis*, 1–28. doi: 10.1214/20-BA1221

Viswesvaran, C., & Ones, D. S. (1995). Theory testing: Combining psychometric meta-analysis and structural equations modeling. *Personnel Psychology*, *48*(4), 865–885. doi: 10.1111/j.1744-6570.1995.tb01784.x

Wu, H., & Browne, M. W. (2015, September). Quantifying adventitious error in a covariance structure as a random effect. *Psychometrika*, *80*(3), 571–600. doi: 10.1007/s11336-015-9451-3

Yao, Y., Vehtari, A., Simpson, D., & Gelman, A. (2018). Using stacking to average Bayesian predictive distributions (with discussion). *Bayesian Analysis*, *13*(3), 917–1007. Retrieved from `https://doi.org/10.1214/17-BA1091` doi: 10.1214/17-BA1091

Yuan, K.-H., & Kano, Y. (2018, December). Meta-analytical SEM: Equivalence between maximum likelihood and generalized least squares. *Journal of Educational and Behavioral Statistics*, *43*(6), 693–720. doi: 10.3102/1076998618787799

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, *67*(6), 361–370.