Assessing local fit by approximating probabilities

R. Noah Padgett

Baylor University

Author Note

R. Noah Padgett, Department of Educational Psychology, Baylor University.

Correspondence concerning this article should be address to R. Noah Padgett,

Department of Educational Psychology, One Bear Place # 97304, Baylor University, Waco,

TX 76798. Contact: `noah_padgett1@baylor.edu`

Abstract

Validity evidence for factor structures underlying a set of items can come from how well a proposed model reconstructs, or fits, the observed relationships. Global model fit is limited in that some components of the proposed model fit better than other components. This limitation has lead to the recommendation of examining fit locally within model components. We describe a new probabilistic approach to assessing local fit using a Bayesian approximation, and illustrate use with a simulated dataset. We show how the posterior approximation closely approximated the sampling distribution of the true parameter. We discuss potential limitation and possible generalizations.

Keywords: local fit, factor analysis, Laplace, posterior probability

Assessing local fit by approximating probabilities

Most methodologists would likely agree that statistical models are an approximation of the process that generated the data we observed. Structural equation models are no exception to this point. The approximation of reality has been described as arising from distributional and structural approximations (Bollen, 2019). Identifying these model misspecifications can be at best difficult and at worst impossible. The task of trying to find evidence for a proposed model is then our focus, and a variety of methods have been developed to help this decision making process.

One popular method of finding evidence for once's proposed model is the use of global fit indices. Global fit indices help identify whether the model specified as a whole *well* approximates the data generating process. Typically used indices are the comparative fit index (Bentler, 1990), RMSEA (Browne & Cudeck, 1992), and SRMR (Bentler, 1995; Jöreskog & Sörbom, 1981; Maydeu-Olivares et al., 2018). Many of these indices are now automatically computed in most SEM software. However, these indices only provide a coarse estimate of model-data fit (Steiger, 2007).

One difficulty comes when the global fit indices are below acceptable (whatever this means to the researcher) and then trying to remedy the misspecification is a systematic and defensible manner. The global fit indices don't necessarily help us identify where the misspecification is occurring, though. Presumably, some parts of the model are better representing the latent structure than other parts, but identifying which parts are not representative is not straightforward. Methods of identifying the source of misfit include investigating residual matrices (Kline, 2015; Maydeu-Olivares & Shi, 2017), modification indicies (Kaplan, 1989; Sörbom, 1989), Wald tests (Buse, 1982; Wald, 1943), likelihood ratio tests (Buse, 1982; Neyman & Pearson, 1928), and model-implied instrumental variables (Bollen, 1995, 2019). Each method has benefits and pitfalls that need to be considered when looking for evidence of model-data fit. The methods above generally provide users with a way of either identifying which specific relationships among observed

that are not capture by the model (e.g., residuals) or which specific component of the model that doesn't fit well (e.g., Wald tests). Once a change in the model has been conducted, a researcher can then compare the results of local fit to see if fit improved.

Evaluating the change in model fit has been helpful for many in the model fit evaluation process. But, this process is not without potential drawbacks. For example, modification indices can lead to conflicting information about sources of misfit leading to statistically equivalent models. Adding too many paths can lead to over fitting to the observed data and difficulty in interpretation. Aside from investigating standardized residuals (Maydeu-Olivares & Shi, 2017), most current methods of investigating local fit rely on non-intuitive metrics (e.g., modification indices, likelihood ratios, or p-values); however, we are proposing a more directly interpretable approach using probabilities for local fit assessment. We aim to develop a probabilistic approach to approximate whether additional paths will result in meaningfully interpretable contribution to the model. With our method, we aim to help control against over fit and aid interpretation driven local fit assessment.

The main contributions of this research are: First, we extend previous research on local fit assessment by introducing a probabilistic method and the describe the conditions in which this approach holds theoretically reliable results. Second, we demonstrate the application of this approach in a simulated dataset with a known latent structure. Finally, we investigated the performance through Monte Carlo simulation study evaluating the accuracy of the predicted probabilities.

Proposed Probabilistic Method

Local fit assessment by means of investigating residual matrices, modification indicies, Wald tests, likelihood ratio tests or model-implied instrumental variables has been shown to provide useful information in a variety of contexts (Chou & Bentler, 1990; Maydeu-Olivares & Shi, 2017; Whittaker, 2012). However, we were still left in a bit of quandary as to whether any potential changes would be of a substantive importance. We

have developed a hopefully straightforward approach to tackling this aspect of local fit evaluation by means of in iterative approximation of the non-estimated parameters. In our experience, we have found that the model changes proposed through modification indices results in parameter estimates that were of little substantive meaning. For example, a cross-loading that is low or a residual covariance that suggests a weak relationship that substantively doesn't add to our understanding of the measurement of the construct of interest.

We propose a method that approximates what the magnitude of the non-estimated parameters would be and to couch the estimates in terms of the probability that the parameter is of substantive interest. The idea is to define a region of the parameter space that we would consider to be of substantive meaningfulness. A natural choice in the case of factor loadings is to define the region as $|\lambda| \geq 0.32$, which is the threshold typically considered as the bound for standardized factor loadings in exploratory factor analysis. However, we are not restricted to this region; for example, we might be interested in knowing if any loadings would likely be greater than 1, indicating that the item loads more strongly than the item we fixed for identification. This approach of approximating regions is similar to defining a "region of practical equivalence (ROPE)" described in Shi et al. (2019) as a Bayesian approach for measurement invariance testing. The ROPE is a region in the parameter space that the researcher determines to be insignificant, and this is already done in most applications of exploratory factor analysis (EFA) when the researcher suppresses factor loadings that are below 0.32 (Benson & Nasser, 1998).

Our idea boils down to defining regions of the parameter space that are of practical insignificant and then approximate the probability that our model parameters fall in that space. The approximation we are proposing is based on a rough approximation of a posterior distribution to compute probabilities based on Bayes theorem. A similar quantity often of interest in Bayesian methods is the posterior predictive probability (PPP or p-value, Gelman et al., 1996; Rubin, 1996). The PPP is often used to construct

probabilistic inferences about test statistics from the posterior predictive distribution. However, for the current application we are interested in computing a probability based on the posterior itself leading to a simpler approximation problem. The proposed local fit assessment builds on this idea of approximating the posterior probability associated the parameters.

A similar approach of assessing model fit was discussed by Lee et al. (2016). The authors used a similar posterior approximation to help evaluate global model fit and identify influential observations. The authors hinted that the probabilistic approach can be used to assess model fit in a variety of ways such as residual covariances. However, to date no one has built on this method for investigations of local fit assessment and is the focus of our work. Next, we introduce the Bayesian ideas needed to build up the approach.

Introduction to Bayesian Approach

Bayes theorem is one of the most powerful tools to manipulate probabilities. Bayes theorem lets us decompose a conditional probability into a likelihood statement about our data and a hypothesized probability space for model parameters (i.e., the prior). Or, more formally as

$$p(\theta|Y) \propto \ell(Y|\theta)p(\theta)$$

where $\propto$ is read as "proportional to", $\ell(Y|\theta)$ is the likelihood of the data Y determined by the model specified and parameters θ, and $p(\theta)$ is the joint prior distribution for the parameters in the model. The posterior is usually sampled from using Monte Carlo methods (e.g., Gibbs sampling, Hamiltonian Monte Carlo, etc.). However, in this case we elect to utilize a *relatively* simple approximation of the posterior that we can sample from relatively quickly.

Under mild regularity conditions the posterior distribution for θ will approximate a normal distribution as the number of samples drawn from the distribution increases (Gelman et al., 2013). Using this property of the posterior, we can use Laplace's method to

derive a second-order Taylor series approximation to the posterior (Tierney & Kadane, 1986). The Taylor series approximation is used to get a point estimate for the posterior expectation and variance. Then, using these two pieces matched with the knowledge that the posterior is asymptotically normally distributed, we arrive at an approximate posterior distribution for any parameter. The approximation is known to provide a useful approximation of posteriors for relatively simple models in a variety of contexts (Kass & Steffey, 1989; Rue et al., 2009).

Under the same regularity conditions that allow us to assume the normal distribution we would get that that $\hat{\theta}$ under the Laplace approximation is nearly equal to $\hat{\theta}_{MLE}$ and that the variance under the Laplace approximation is nearly identical to the I_{OBS} (the observed information). Not including the prior distribution for θ has been done previously (Lee et al., 2016; Li, 1992; Rue et al., 2009; Wolfinger, 1993). This would decrease our uncertainty in the posterior distribution for θ even if only slightly. Additionally, utilizing a minimally informative prior structure provides a useful initializing point for numerical methods. Next, we will describe how we can utilize Laplace's method in factor analysis for local fit assessment.

Applying Laplace Approximation to CFA

We borrowed most notation for Bayesian factor analysis from the chapter (). We can take the traditional factor model (shown below) and restructure the model to be able to efficiently approximate posterior features of interest. Full model is

$$\mathbf{Y}_i = \tau + \Lambda\eta_i + \theta_i$$

$$\Sigma = \Lambda^T \Phi \Lambda + \Theta$$

$$\mathbf{Y}_i \mid \tau, \Lambda, \eta_i, \Theta \sim \mathrm{N}\left(\tau + \Lambda\eta_i, \Theta\right)$$

$$\eta_i \mid \kappa, \Phi \sim \mathrm{N}\left(\kappa, \Phi\right)$$

where $\mathbf{Y}_i$ is the response vector for person i, τ is the vector of indicator intercepts, Λ is the matrix of factor loadings, η_i is the vector of factor scores, θ_i is the vector of residuals, Σ is

the indicator/observed variable covariance matrix, Θ is the error covariance matrix, κ is the vector of factor intercepts (usually fixed to 0), and Φ is the factor covariance matrix. Which is a pretty large model with a lot of potential parameters and moving pieces. The full model could be incorporated in a Bayesian perspective relatively easily, but the parameter space grows rapidly. Below is one potential way the full joint posterior could be express.

$$\pi\left(\tau, \Lambda, \eta, \Theta, \kappa, \Phi \mid Y\right) \propto \ell\left(Y \mid \tau, \Lambda, \eta, \Theta, \kappa, \Phi\right) \times \pi\left(\tau, \Lambda, \eta, \Theta, \kappa, \Phi\right)$$

Which is pretty gnarly if there are a large number of subjects, items, and latent variables. The likelihood is not difficult to define due to the assumption of conditional independence of subjects, and even further simplification can be made if conditional independence of items is assumed (i.e., Θ is a diagonal matrix)

$$\ell\left(Y \mid \tau, \Lambda, \eta, \Theta, \kappa, \Phi\right) = \prod_{i=1}^{N} f\left(\mathbf{Y}_i \mid \tau, \Lambda, \eta_i, \Theta\right)$$
$$= \prod_{i=1}^{N} \prod_{j=1}^{J} f\left(Y_{ij} \mid \tau_j, \lambda_j, \eta_i, \theta_{jj}\right)$$

which results in a nice linerization of the complex looking likelihood that we started with. But, the more complicated issue is on the construction of the joint prior distribution. Constructing joint prior distributions directly is typically not possible outside of special cases (e.g, normal-gamma conjugate prior in normal theory multiple linear regression) in multiparameter problems.

In multiparameter problems, a common method for building the join prior distribution is to construct the prior hierarchically. The hierarchical prior for factor models can be constructed by decomposing the prior into parameter groups that can defensibly be independent or at least conditionally independent. A potential decomposition is to have independent groups for factor loadings (Λ), item intercepts (τ), and error (co)variances (Θ), then group the factor scores (η), factor intercepts (κ), and factor covariance matrix (Φ) together. The later grouping can be structured where the factor scores distribution is conditional on the factor intercept and factor covariance matrix priors. The decomposed

prior could look something like this

$$\pi\left(\tau,\Lambda,\eta,\Theta,\kappa,\Phi\right)=\pi\left(\eta,\kappa,\Phi\right)\pi\left(\tau\right)\pi\left(\Lambda\right)\pi\left(\Theta\right)$$

$$=\pi\left(\eta\mid\kappa,\Phi\right)\pi\left(\kappa\right)\pi\left(\Phi\right)\pi\left(\tau\right)\pi\left(\Lambda\right)\pi\left(\Theta\right).$$

Some common specifications of the priors for these parameters are

$$\eta\mid\kappa,\Phi\sim\mathrm{N}(\kappa,\Phi)$$

$$\kappa\sim\mathrm{N}(\mu_\kappa,\sigma_\kappa^2)$$

$$\Phi\sim\mathrm{Inverse-Wishart}(d\Phi_0,d),\ d\geq M$$

$$\tau j\sim\mathrm{N}(\mu_\tau,\sigma_\tau^2)$$

$$\lambda_{jm}\sim\mathrm{N}(\mu_\lambda,\sigma_\lambda^2)$$

$$\theta_{jj}\sim\mathrm{Inverse-Gamma}(\alpha_\theta,\beta_\theta).$$

The indices j correspond to the jth indicator and m is index over latent variables. The factor loadings, intercepts, error (co)variances may be specified independently for each individual parameter which greatly simplifies the process, but is not necessary. If one had a belief that the error (co)variance matrix was not diagonal, then that structure could be included as part of a joint prior on Θ.

Once the joint prior is constructed, one difficult hurdle of a full Bayesian analysis would be to try to sample from the posterior. This is generally where we would move to some type of Markov Chain Monte Carlo set-up to try to efficiently sample from the marginal posteriors of these parameters. Some software options that make this job easier include M*plus* (Muthén & Muthén, 2017), blavaan (Merkle & Rosseel, 2018), JAGS (Plummer et al., 2003), OpenBUGS (Surhone et al., 2010), Stan (Carpenter et al., 2017), and many more. However, in this study, we are proposing a drastic simplification of this process to aim a a rough approximation of specific features of the posterior.

As we have described in the section "Proposed Probabilistic Method", our aim is to approximate the probability that a parameter is outside a region of practical importance (or that the absolute value exceeds a threshold). The powerful machinery that is MCMC can

certainly be used for this purpose. However, our intent is to provide a simpler, quicker, and less computationally intensive approach that can be utilized as part of the model evaluation process for local fit assessment. The complex posterior distribution built up till now is simplified as we shift from sampling from the marginal posterior to a specific conditional posterior for each parameter. Meaning we sample from each posterior by fixing estimates of loadings, error (co)variances, factor (co)variances, etc. in order to reduce the model down to a single parameter problem. Then a very simple and generalize-able method is available for approximately sampling from the conditional posterior of parameter(s) of interest.

This approach may work as follows for approximating the posterior of a specific factor loading λ_{z1}, the zth loading of factor 1. If we utilize the nice simplifications likelihood that result from the independence assumptions, then the posterior simplifies to the

following by conditioning on the fixed parameters that can come from the MLE solution.

$$\pi(\lambda_{z1} \mid \theta) = \pi(\lambda_{z1} \mid \tau, \Lambda_{zm \neq z1}, \eta, \Theta, \kappa, \Phi, \mathrm{Y})$$

$$\propto \ell\left(\mathrm{Y} \mid \tau, \Lambda, \eta, \Theta, \kappa, \Phi\right) \times \pi(\lambda_{z1})$$

$$= \prod_{i=1}^{N} f\left(\mathbf{Y}_i \mid \tau, \Lambda, \eta_i, \Theta\right) \times \pi(\lambda_{z1} \mid \mu_\lambda, \sigma_\lambda^2)$$

$$= \prod_{i=1}^{N} \prod_{j=1}^{J} f\left(Y_{ij} \mid \tau_j, \lambda_j, \eta_i, \theta_{jj}\right) \times \pi(\lambda_{z1} \mid \mu_\lambda, \sigma_\lambda^2)$$

$$\propto \prod_{i=1}^{N} f\left(Y_{iz} \mid \tau_z, \lambda_{z1}, \eta_{i1}, \theta_{zz}\right) \times \pi(\lambda_{z1} \mid \mu_\lambda, \sigma_\lambda^2)$$

$$= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_{zz}^2}} \exp\left\{\frac{-1}{2\sigma_{zz}}\left(Y_{iz} - (\tau_z + \lambda_{z1}\eta_{i1})\right)^2\right\} \frac{1}{\sqrt{2\pi\sigma_\lambda^2}} \exp\left\{\frac{-1}{2\sigma_\lambda}\left(\lambda_{z1} - \mu_\lambda\right)^2\right\}$$

$$\propto \exp\left\{\frac{-(\sigma_{zz} + \sigma_\lambda \sum \eta_{i1}^2)}{2\sigma_\lambda \sigma_{zz}}\left(\lambda_{z1} - \frac{\sigma_\lambda \sum y_{iz}\eta_{i1} - \sigma_\lambda \tau_z \sum \eta_{i1} + \sigma_{zz}\mu_\lambda}{\sigma_{zz} + \sigma_\lambda \sum \eta_{i1}^2}\right)^2\right\}$$

$$\therefore$$

$$\lambda_{z1} \sim \mathrm{N}(M, V)$$

$$M = \frac{\sigma_\lambda \sum y_{iz}\eta_{i1} - \sigma_\lambda \tau_z \sum \eta_{i1} + \sigma_{zz}\mu_\lambda}{\sigma_{zz} + \sigma_\lambda \sum \eta_{i1}^2}$$

$$V = \left(\frac{\sigma_\lambda \sigma_{zz}}{\sigma_{zz} + \sigma_\lambda \sum \eta_{i1}^2}\right)^2$$

which can be reduced more depending on the of priors for the λ_{z1}. But, this is not the most helpful given that the posterior depends on the unknown values of the latent variable η, but thankfully in application we can side-step the need for these values. Essentially, what we have done above is just a straightforward application of Bayes theorem to get a closed form solution for the posterior. This is not as simple for the residual covariances, though. We think the above was mostly just a sanity check to make sure we knew what the math was doing and to show that under the simplifications we're proposing we can arrive at "nice" solutions. But, this is certainly not going to be done in practice given that the posteriors for cross-loadings and (co)variances are typically more difficult to derive depending on the choice of prior.

Illustrative Example

So, now to illustrate the method, suppose we have the "true" factor model shown in Figure 1. A sample of 300 was drawn from this population. We simplified the model here for illustration of our approach to local fit assessment. Using this sample, a factor analysis
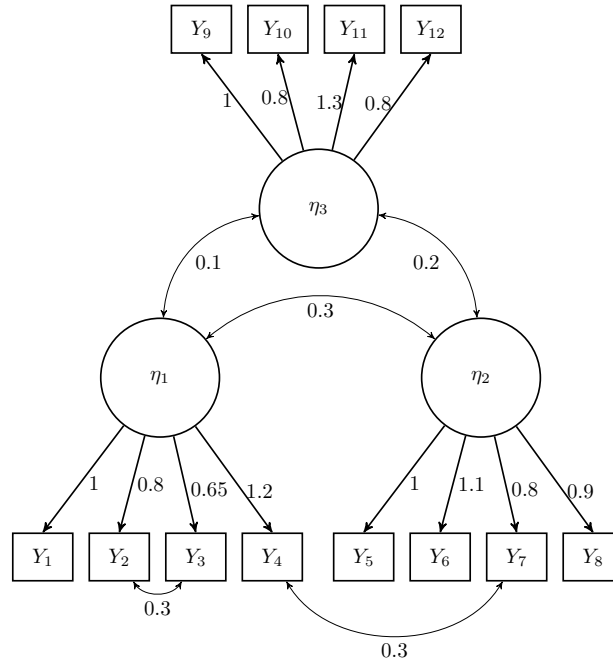

Figure 1
Simulating data model

without the residual covariances was fit to these data using lavaan (Rosseel, 2012). We found evidence that the model does not fit these data ($\chi^2(51) = 72.8, p = 0.024$). Now the task is to try to find out which relationships we are missing or have underestimated according to our model. Next, we applied the described probabilistic approach to try to identify likely omitted paths that are the source of misfit. These results are shown next whether the paths identified as most likely are shown in descending order. For simplicity we are only showing the paths that had probabilities greater than 0.10 (i.e., 100 out of 1000 draws from the posterior were outside the region of practical equivalence). These estimates are shown in Table 1. We found that, with the method described in this paper, the two paths with the highest probability of being meaningful in magnitude are the two

residual covariances that were in the population model.

Table 1

Estimated probabilities of parameters being outside region of practical equivalence

Parameter (θ)	$\Pr(\mid \theta \mid \geq cutoff)$
$\mathrm{cov}(y_7, y_4)$	0.810
$\mathrm{cov}(y_3, y_2)$	0.461
$\mathrm{cov}(y_{12}, y_7)$	0.254
$\mathrm{cov}(y_9, y_4)$	0.233
$\mathrm{cov}(y_4, y_2)$	0.221
$\mathrm{cov}(y_{10}, y_1)$	0.201
$\mathrm{cov}(y_9, y_1)$	0.191
$\mathrm{cov}(y_9, y_7)$	0.164
$\mathrm{cov}(y_9, y_8)$	0.161
$\mathrm{cov}(y_4, y_1)$	0.160
$\mathrm{cov}(y_3, y_1)$	0.147
$\mathrm{cov}(y_9, y_2)$	0.126
$\mathrm{cov}(y_{10}, y_5)$	0.108

Simulation Study

In this simulation study, we evaluated the estimated probabilities against the sampling distribution of the parameters. To do so, we generated 10,000 datasets from the population model, fit the population model to each, and extracted the sampling distribution of the model parameters. We compared the sampling distribution to the approximated empirical sampling distributions used to compute the approximating probabilities shown in Table 1. We focused on the two residual covariances from the model in Figure 1, because these two parameters are likely to be initially omitted from a model specification.

The true and approximated sampling distributions are shown in Figure 2. The probabilities from the "true" sampling distribution are $\Pr(\mathrm{cov}(y_7, y_4) > 0.25) = .82$ and $\Pr(\mathrm{cov}(y_3, y_2) > 0.25) = .72$. We found relatively major difference between the true and approximated sampling distribution for $\mathrm{cov}(y_3, y_2)$ and the probability. This is in contrast to the nearly perfectly matched distribution for $\mathrm{cov}(y_7, y_4)$.
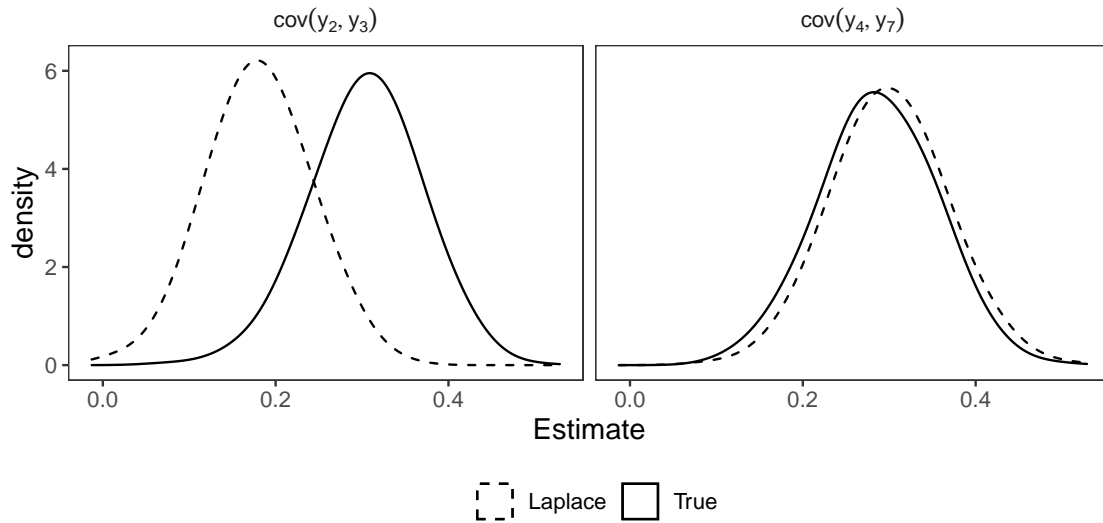
Figure 2

True and approximated sampling distributions for model residual covariances

Discussion

In this paper, we have outlined our method investigating local fit by approximating the probability that a parameter is meaningfully different than zero. We built our approach on similar ideas of model fit by Lee et al. (2016) and Shi et al. (2019). By utilizing a Bayesian perspective, the resulting probability is intuitive to interpret. A result that doesn't require a complex interpretation provides a nice compliment to existing methods of model modification (e.g., Kaplan, 1989; Sörbom, 1989; Wald, 1943). We avoided some of the complexities of estimation by using a straightforward approximation based on Laplace's method. The approximation is also relatively fast compared to full Markov chain Monte Carlo methods.

Despite some potential advantages of the approach we have developed, we are also aware of some potential disadvantages as well. One of the issues is trying to decide what value of each parameter should be used for computing the probability? For (standardized) factor loadings, we used a value of 0.32 suggest by Benson and Nasser (1998) from EFA, but other values may be of interest. More work needs to be done to figure out how strong correlations need to be in order to be of interest. However, these values probably need to

change depending on the inferences of interest. Additionally, we have a similar issue of trying to decide how high does the probability need to be in order for us to have confidence that the omitted path is meaningful? 0.5? 0.4? Maybe this value should be a function of the number of parameters tested, similar to a Bonferroni adjustment. A set number for this probability may be counter productive since we want to view this probability more as reflective of degree of evidence or belief about the magnitude of the parameter. We view a higher probability is more evidence that the parameter is of interest. This is sometimes called the *epistemic* or *degree-of-belief* perspective of probability (de Finetti, 1974; Levy & Mislevy, 2016, p.XX).

We have provided an approach to local fit assessment that allows for more information to be gained about the relationships in our data. Our proposed approach is intuitive and aligns with aims to search for sources of local fit rather than global fit. Future work could investigate how this approach could be utilized for parameters included in the model not just the omitted paths.

References

Benson, J., & Nasser, F. (1998). On the Use of Factor Analysis as a Research Tool. *Journal of Vocational Education Research*, *23*(1), 13.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, *107*(2), 238–246. https://doi.org/10.1037/0033-2909.107.2.238

Bentler, P. M. (1995). *EQS structural equations program manual*. Multivariate Software.

Bollen, K. A. (1995). Structural equation models that are nonlinear in latent variables: A least-squares estimator. *Sociological Methodology*, *25*, 223. https://doi.org/10.2307/271068

Bollen, K. A. (2019). Model Implied Instrumental Variables (MIIVs): An Alternative Orientation to Structural Equation Modeling. *Multivariate Behavioral Research*, *54*(1), 31–46. https://doi.org/10.1080/00273171.2018.1483224

Browne, M. W., & Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, *21*(2), 230–258. https://doi.org/10.1177/0049124192021002005

Buse, A. (1982). The Likelihood Ratio, Wald, and Lagrange Multiplier Tests: An Expository Note. *The American Statistician*, *36*(3a), 153–157. https://doi.org/10.1080/00031305.1982.10482817

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Chou, C. P., & Bentler, P. M. (1990). Model Modification in Covariance Structure Modeling: A Comparison among Likelihood Ratio, Lagrange Multiplier, and Wald Tests. *Multivariate Behavioral Research*, *25*(1), 115–136. https://doi.org/10.1207/s15327906mbr2501__13

de Finetti, B. (1974). *Theory of probability, violumne 1*. Wiley.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd). CRC press.

Gelman, A., Meng, X. L., & Stern, H. (1996). Posterior predictive assessment of model fitness via realized discrepancies. *Statistica Sinica*, *6*(4), 733–807.

Jöreskog, K. G., & Sörbom, D. (1981). *LISREL V: Analysis of linear structural relationships by the method of maximum likelihood*. National Educational Resources Jöreskog, K. G., & Sörbom, D. (1981). USREL V: Analysis of linear structural relationships by the method of maximum likelihood. Chicago: National Educational Resources.

Kaplan, D. (1989). Model Modification in Covariance Structure Analysis: Application of the Expected Parameter Change Statistic. *Multivariate Behavioral Research*, *24*(3), 285–305. https://doi.org/10.1207/s15327906mbr2403_2

Kass, R. E., & Steffey, D. (1989). Approximate bayesian inference in conditionally independent hierarchical models (Parametric empirical bayes models). *Journal of the American Statistical Association*, *84*(407), 717–726. https://doi.org/10.1080/01621459.1989.10478825

Kline, R. B. (2015). *Principles and practice of structural equation modeling* (Fourth). The Guilford Press.

Lee, T., Cai, L., & Kuhfeld, M. (2016). A poor person's posterior predictive checking of structural equation models. *Structural Equation Modeling*, *23*(2), 206–220. https://doi.org/10.1080/10705511.2015.1014041

Levy, R., & Mislevy, R. J. (2016). *Bayesian psychometric modeling*. CRC Press.

Li, B. (1992). Laplace expansion for posterior densities of nonlinear functions of parameters. *Biometrika*, *79*(2), 393–401.

Maydeu-Olivares, A., & Shi, D. (2017). Effect sizes of model misfit in structural equation models: Standardized residual covariances and residual correlations. *Methodology*, *13*(Supplement 1), 23–30. https://doi.org/10.1027/1614-2241/a000129

Maydeu-Olivares, A., Shi, D., & Rosseel, Y. (2018). Assessing Fit in Structural Equation

    Models: A Monte-Carlo Evaluation of RMSEA Versus SRMR Confidence Intervals

    and Tests of Close Fit. *Structural Equation Modeling*, *25*(3), 389–402.

    https://doi.org/10.1080/10705511.2017.1389611

Merkle, E. C., & Rosseel, Y. (2018). blavaan: Bayesian structural equation models via

    parameter expansion. *Journal of Statistical Software*, *85*(4), 1–30.

    https://doi.org/10.18637/jss.v085.i04

Muthén, L., & Muthén, B. (2017). *Mplus User's Guide* (8th). Muthén & Muthén.

Neyman, J., & Pearson, E. S. (1928). On the use and interpretation of certain test criteria

    for purposes of statistical inference. *Biometrika*, *20*(3/4), 175–240, 263–294.

    https://doi.org/10.2307/2332112

Plummer, M. et al. (2003). Jags: A program for analysis of bayesian graphical models using

    gibbs sampling. *Proceedings of the 3rd international workshop on distributed

    statistical computing*, *124*(125.10), 1–10.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more.

    Journal of statistical software, *48*(2), 1–36.

Rubin, D. B. (1996). Comment: On posterior predictive p-values. *Statistica Sinica*, *6*(4),

    787–792.

Rue, H., Martino, S., & Chopin, N. (2009). Approximate Bayesian inference for latent

    Gaussian models by using integrated nested Laplace approximations. *Journal of the

    Royal Statistical Society. Series B: Statistical Methodology*, *71*(2), 319–392.

    https://doi.org/10.1111/j.1467-9868.2008.00700.x

Shi, D., Song, H., DiStefano, C., Maydeu-Olivares, A., McDaniel, H. L., & Jiang, Z. (2019).

    Evaluating Factorial Invariance: An Interval Estimation Approach Using Bayesian

    Structural Equation Modeling. *Multivariate Behavioral Research*, *54*(2), 224–245.

    https://doi.org/10.1080/00273171.2018.1514484

Sörbom, D. (1989). Model modification. *Psychometrika*, *54*(3), 371–384.

    https://doi.org/10.1007/BF02294623

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural

    equation modeling. *Personality and Individual Differences*, *42*(5), 893–898.

    https://doi.org/10.1016/j.paid.2006.09.017

Surhone, L. M., Tennoe, M. T., & Henssonow, S. F. (2010). *Openbugs*. Betascript

    Publishing. https://doi.org/10.5555/1941071

Tierney, L., & Kadane, J. B. (1986). Accurate approximations for posterior moments and

    marginal densities. *Journal of the American Statistical Association*, *81*(393), 82–86.

    https://doi.org/10.1080/01621459.1986.10478240

Wald, A. (1943). Tests of Statistical Hypotheses Concerning Several Parameters When the

    Number of Observations is Large. *Transactions of the American Mathematical*

    Society, *54*(3), 426. https://doi.org/10.2307/1990256

Whittaker, T. A. (2012). Using the Modification Index and Standardized Expected

    Parameter Change for Model Modification. *The Journal of Experimental Education*,

    80(1), 26–44. https://doi.org/10.1080/00220973.2010.531299

Wolfinger, R. (1993). Laplace's approximation for nonlinear mixed models. *Biometrika*,

    80(4), 791–796.