

CHAPTER THREE

**Alternatives to Traditional Model Comparison
Strategies for Covariance Structure Models**

Kristopher J. Preacher

University of Kansas

Li Cai

Robert C. MacCallum

University of North Carolina at Chapel Hill

APPENDIX

One of the most useful formulations of MDL for model selection purposes is termed the *normalized maximum likelihood* (NML; Rissanen, 2001):

$$\text{NML} = \frac{L(y|\hat{\theta})}{\int_S L(z|\hat{\theta}(z)) dz} \quad (\text{A1})$$

NML represents the information in a model (i.e., the likelihood with respect to given data y in the numerator) adjusted for the model’s general data-fitting capabilities. This adjustment is located in the denominator, where the multiple integral is a rescaling of the mean likelihood when the model is fit to all possible data z . The denominator also acts as a normalizing constant for the numerator. Because integration is taken over the data space (S) and not the parameter space, this expression of NML is invariant under model reparameterization.

Because computation of the denominator term of NML is intractable and approximation is difficult in many applications, it is sometimes more expedient to use an earlier formulation of MDL, the *Fisher information approximation* (FIA):

$$\text{FIA} = -\ln L(y|\hat{\theta}) + \frac{q}{2} \ln \left(\frac{N}{2\pi} \right) + \ln \int_{\Theta} \sqrt{|I(\theta)|} d\theta \quad (\text{A2})$$

where $|I(\theta)|$ is the determinant of the Fisher information matrix. FIA converges asymptotically to $-\ln(\text{NML})$ (Rissanen, 1996), and thus can be used to closely approximate the rank ordering of models that would be achieved under NML. The FIA expression of the MDL principle has the benefit of explicitly illustrating the sources of model complexity. The contribution of the number of free parameters (q) to complexity, weighted by sample size, is explicit in the second term. The third term, a multiple integral taken over the parameter space (Θ), represents structural complexity. As parameters become more redundant, the determinant of $I(\theta)$ decreases, leading to a smaller integral. Thus, as the determinant increases, FIA becomes larger, essentially imposing a penalty due to model structure. Furthermore, in order to use FIA, all model parameters must be identified. Otherwise, $I(\theta)$ becomes rank-deficient, leading to infinite values of $\int_{\Theta} \sqrt{|I(\theta)|} d\theta$. For models with more than two or three parameters, $\ln \int_{\Theta} \sqrt{|I(\theta)|} d\theta$ can be just as difficult to compute as $\int_S L(z|\hat{\theta}(z)) dz$, so numerical methods are normally employed (Zhang, 1999). A useful approximation to FIA (Hansen & Yu, 2001; Markon & Krueger, 2004; Rissanen, 1989) is *stochastic information complexity* (SIC):

$$\text{SIC} = -\ln L(y|\hat{\theta}) + \frac{1}{2} \ln |nI(\theta)|, \tag{A3}$$

which is clearly more tractable in applied settings, assuming the information matrix is accessible. All three of the MDL expressions presented here require that certain regularity conditions apply. FIA and SIC require that the likelihood function have a unique maximum and be twice differentiable (Lanternman, 2001). NML requires only that a unique maximum exist and that the integral converge.