The Issue of Isopower in Power Analysis for Tests of Structural Equation Models

Robert MacCallum , Taehun Lee & Michael W. Browne

Ψ Psychology Press
Taylor & Francis Group

The Issue of Isopower in Power Analysis for Tests of Structural Equation Models

Robert MacCallum and Taehun Lee
University of North Carolina at Chapel Hill

Michael W. Browne
The Ohio State University

Two general frameworks have been proposed for evaluating statistical power of tests of model fit in structural equation modeling (SEM). Under the Satorra–Saris (1985) approach, to evaluate the power of the test of fit of Model A, a Model B, within which A is nested, is specified as the alternative hypothesis and considered as the true model. We then determine the power of the test of fit of A when B is true. Under the MacCallum–Browne–Sugawara (1996) approach, power is evaluated with respect to the test of fit of Model A against an alternative hypothesis specifying a true degree of model misfit. We then determine the power of the test of fit of A when a specified degree of misfit is assumed to exist as the alternative hypothesis. In both approaches the phenomenon of isopower is present, which means that different alternative hypotheses (in the Satorra–Saris approach) or combinations of alternative hypotheses and other factors (in the MacCallum–Browne–Sugawara approach) yield the same level of power. We show how these isopower alternatives can be defined and identified in both frameworks, and we discuss implications of isopower for understanding the results of power analysis in applications of SEM.

Statistical power analysis is widely used to aid in research planning across a range of research designs and statistical models and methods. A generic power analysis involves several steps, beginning with establishment of a null hypothesis, H_0, to be tested, where H_0 usually specifies a testable hypothesis about a condition in the population of interest. Typical null hypotheses specify values of one or more population parameters or functions of parameters. An alternative hypothesis, H_1, is specified and is assumed to be true for purposes of power analysis, with the difference between H_1 and H_0 representing effect size, or the degree to which H_0 is false. Given the distribution of a test statistic under H_0 and under H_1, it is then possible to determine statistical power, or the probability of rejecting H_0 when H_1 is true. The outcome of such an analysis depends on various factors including effect size, sample size, and the α level to be

used in the test of H_0. Power analysis provides the investigator with a capability for estimating the likelihood of detecting an anticipated effect, or an effect of scientific or practical interest. It also provides for determination of the size of an effect that can be detected with high likelihood given other design characteristics. Power analysis is also commonly used to determine minimum sample size (N) necessary to detect a particular effect with a desired level of power.

The phenomenon of *isopower* involves the existence of different alternative hypotheses (possibly in combination with changes in other factors) that yield the same level of power for a test of a given null hypothesis. Implicitly this phenomenon means that the same effect size might be produced by different H_1s. Thus the outcome of a given power analysis might not be uniquely a function of the particular H_1 employed in a computation of statistical power, which means that the particular H_1 might more appropriately be considered as a member of a class of *isopower alternatives,* with all H_1s in that class yielding the same power. In practice, if such isopower alternatives exist, the investigator may benefit from identifying at least some of them so as to gain more information about effects associated with a given level of power. Such information might better inform research design and interpretation of results. These issues are discussed further later in this article.

We examine here the issue of isopower in the context of power analysis for tests of structural equation models. Given a structural equation model and a set of sample data to which the model can be fit, and assuming the use of maximum likelihood (ML) estimation, the model of interest is typically tested using the likelihood ratio test. The conventional H_0 is that the model is correct in the population, or that the population covariance matrix, $\mathbf{\Sigma}$, is exactly explained by the model. Equivalently we can specify the null hypothesis as H_0: $F_0 = 0$, where F_0 is the population value of the ML discrepancy function used for estimation, and the alternative hypothesis is H_1: $F_0 > 0$. Thus, H_0 represents perfect fit in the population and H_1 represents imperfect fit. The null hypothesis is tested using the likelihood ratio test statistic, $T = (N - 1)\hat{F}$, where $\hat{F}$ is the sample value of the ML discrepancy function. Under H_0 and multivariate normality T is asymptotically distributed as χ^2 with degrees of freedom $d = p(p + 1)/2 - q$, where p is the number of measured variables in the model and q is the effective number of parameters estimated. The H_1 in the usual test of H_0 is that $\mathbf{\Sigma}$ is an arbitrary unstructured covariance matrix. The sample χ^2 is compared with a critical χ^2 and a significant result indicates rejection of the H_0 of perfect model fit in the population, whereas a nonsignificant result indicates that H_0 is statistically plausible.

Investigators evaluating structural equation models are often interested in the statistical power of the likelihood ratio test just described. However, power analysis for this test cannot be conducted in exactly this framework because the alternative hypothesis, H_1: $F_0 > 0$, is not sufficiently well defined to allow us to completely specify the distribution of T under H_1. A more specific H_1, specifying a particular value of F_0, is needed to allow for defining that distribution. Two different approaches have been developed for establishing useful H_1s for power analysis in this context, one originating in work by Satorra and Saris (1985) and the other originating with MacCallum, Browne, and Sugawara (1996). These two frameworks use the same body of statistical theory for determining power, but very different concepts and procedures for establishing H_1. In this article we review each of these approaches briefly and, for each, we then closely examine the phenomenon of isopower. We show how isopower alternatives can be defined and discuss the implications of their existence for power analysis for tests of structural equation models.
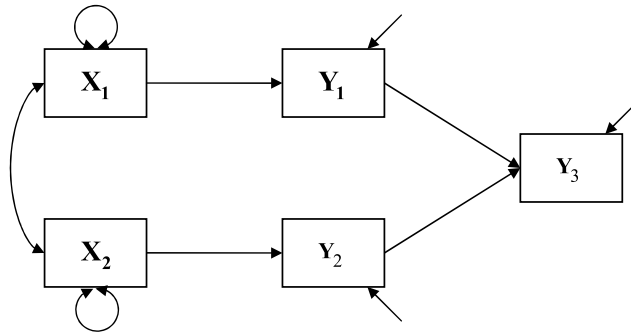
FIGURE 1   Model A representing null hypothesis H_0.

ISOPOWER IN THE SATORRA–SARIS FRAMEWORK

Defining H_1 in Terms of Parametric Misspecification

Suppose the model in Figure 1 represents a Model A to be tested and that the investigator wishes to carry out power analysis for the likelihood ratio test of this model. This process requires specification of an alternative hypothesis, H_1. Satorra and Saris (1985) defined an H_1 by specifying an alternative Model B, where A is nested in B. For illustrative purposes suppose the Model B in Figure 2 represents this H_1. An important requirement of this approach is that the investigator must also specify values of all of the parameters in B. Again, Figure 2 shows such numerical values for purposes of illustration. These values might be based on prior research or might simply represent plausible values or values of interest to the investigator. For purposes of power analysis, Model B is assumed to hold exactly in the population. As shall be seen, power analysis then provides the probability of rejecting A when B, with its specified parameter values, is true.

Of particular importance are the values of the parameters that differentiate A and B, which are designated β_1 and β_2 in Figure 2. The assigned values of these parameters establish a kind
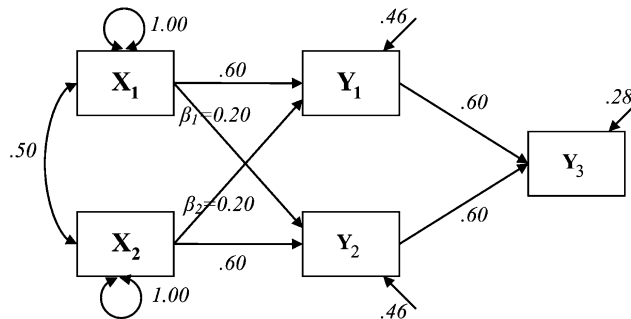

FIGURE 2   Model B representing alternative hypothesis H_1.

of effect size—the degree to which Model A is misspecified, or in other words, the degree to which H_0 is false. Note that under H_0 the parameters β_1 and β_2 have hypothesized values of zero, but under H_1 they are assigned values of .20. Thus, H_1 defines particular parametric misspecifications of the model being tested, Model A.

Following the approach described by Satorra and Saris (1985) the actual power computation proceeds as follows. From Model B with its assigned parameter values we compute the implied covariance matrix, Σ_B. Model A is then fit to Σ_B by ML estimation, with N specified as the N of interest for the power analysis. This step yields a discrepancy function value, $F_{A|B}$ reflecting lack of fit of A when B is true. Under this H_1 the test statistic, T, follows a noncentral chi-square distribution with degrees of freedom $d = p(p + 1)/2 - q$ and noncentrality parameter $\lambda_{A|B} = (N - 1)F_{A|B}$. Power is then computed in the usual fashion, as illustrated in Figure 3. The distribution on the left in Figure 3 represents the distribution of T under H_0, which is central chi-square with $d = p(p + 1)/2 - q$, and the distribution on the right represents the distribution of T under H_1, which is noncentral chi-square with the same degrees of freedom, and noncentrality parameter $\lambda_{A|B} = (N - 1)F_{A|B}$. A critical value of chi-square is determined on the distribution of T under H_0, and power is then computed as the area beyond that critical value under the distribution of T under H_1.

In the illustration based on the models in Figures 1 and 2, when we fit A to Σ_B we find $F_{A|B} = .126$. If we set $N = 100$ we find $\lambda_{A|B} = 12.50$. Both distributions have 5 degrees of freedom. The power computation under these conditions yields a value of .79, meaning that if B is true in the population with its assigned parameter values, and if we carry out the likelihood ratio test of Model A using a random sample of $N = 100$ and $\alpha = .05$, the probability of rejecting the misspecified Model A would be .79.
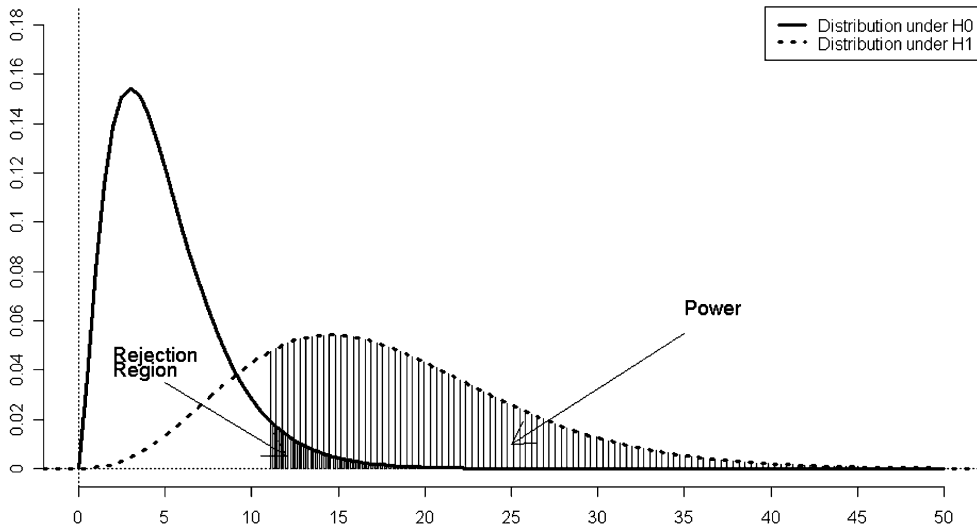

FIGURE 3   Null and alternative distributions of the test statistic for determining statistical power.

Isopower Under Parametric Misspecification of the Tested Model

Given this framework for specification of H_1, the isopower issue then involves the possible existence of other misspecified models that would yield the same level of power. We show that there generally exists an infinite number of such alternatives. We consider first how this phenomenon could occur generically. Note that the distribution of the test statistic under H_0 is fixed, and that the distribution of T under H_1, which establishes the degree of misspecification of A and in turn the value of the noncentrality parameter for the distribution of T under H_1, will determine the level of power. Thus, any alternative model, or H_1, that yields the same noncentral distribution of T will result in the same level of power for the test of H_0. More specifically, because the critical aspect of that distribution under H_1 is the noncentrality parameter, $\lambda_{A|B}$, and because $\lambda_{A|B} = (N-1)F_{A|B}$, then any Model B that yields the same $F_{A|B}$ when A is fit to Σ_B will result in the same power for the test of Model A. The set of all such alternative models that have this property defines a class of isopower alternatives. We consider next several different ways to identify and generate such alternatives.

Case 1: Alternative Values of Parameters for Model B

In the example represented in Figure 2, the parameters labeled β_1 and β_2 are assigned values of .20 and the resulting power of the test of exact fit of Model A (in Figure 1) is .79. Saris and Satorra (1993) pointed out that different values of β_1 and β_2 could be determined that would also yield power of .79 for the test of Model A. In fact, an infinite number of pairs of such values would exist, and Saris and Satorra presented a computational method for determining an "isopower contour" that represents all such possible values. Following their procedure we have obtained the isopower contour shown in Figure 4 for our running example. Each point on the ellipse in Figure 4 represents a pair of values of β_1 and β_2 that, when combined with
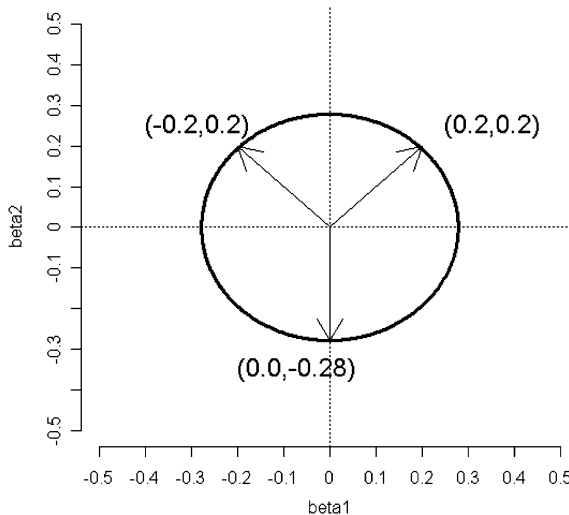

FIGURE 4   Pairs of parameter values (β_1, β_2) producing power of .79.

the fixed values of the other parameters in Model B in Figure 2, would yield power of .79 for the test of Model A. For example, if we assign $\beta_1 = -0.2$ and $\beta_2 = +0.2$, and hold other parameters at their values shown in Figure 2, then the test of Model A when Model B is true would have power of .79. Another isopower alternative is obtained by setting $\beta_1 = 0.0$ and $\beta_2 = -0.28$.

As the number of parameters that differentiate Model B from A increases, this isopower contour would become a hyperellipse in a higher dimensional space, but the general principle still holds. That is, there are alternative values for the parameters that differentiate B from A that yield the same level of power for the test of A. (Note that if B is differentiated from A by only one parameter, β_1, and a value is assigned to β_1 for purposes of power analysis, then this type of isopower analysis will yield only one alternative isopower value for β_1; that is, there would be two isopower points rather than an isopower contour.)

Case 2: Equivalent Models

It has been shown that for virtually any structural equation model it is possible to construct differently parameterized models that will fit any data exactly as well as the original model (Lee & Hershberger, 1990; MacCallum, Wegener, Uchino, & Fabrigar, 1993; Stelzl, 1986). In the present context, this phenomenon can be seen to generate alternative models that represent isopower alternatives. For example, any model that is equivalent to Model B in Figure 2 will yield the same power when used as an alternative in the test of Model A. Given Model B, one could implement rules presented by Lee and Hershberger (1990) to generate equivalent models. For instance, this approach could produce Model B1 in Figure 5. The issue in power analysis then is to define parameter values in B1 that would yield the same implied covariance matrix as does B. This could be achieved easily by constructing Σ_B from the parameter values in Model B (Figure 2), then fitting Model B1 (Figure 5) to that Σ_B. The parameter values in the resulting solution establish the full alternative hypothesis; those values are shown in Figure 5. The covariance matrix Σ_{B1} implied by Figure 5 is identical to Σ_B implied by Figure 2. Obviously, when Model A is fit to Σ_{B1} and power analysis computations are completed, the resulting power would be .79, the same as when Model B is used as the alternative hypothesis.
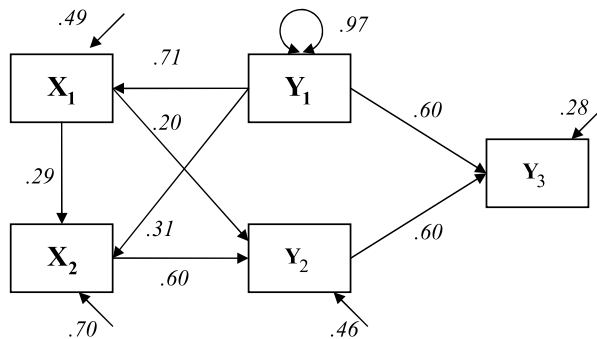

FIGURE 5   Model B1: An isopower alternative produced using equivalent models.
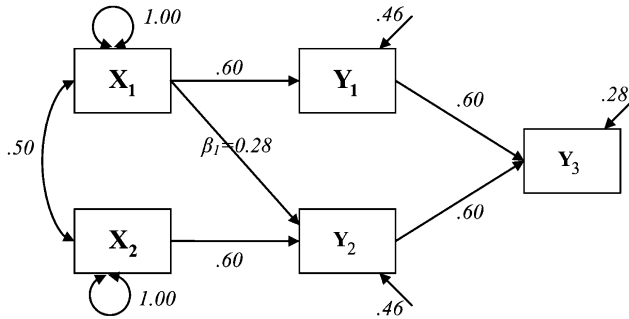
FIGURE 6   Model B2: An isopower alternative in which Model A is nested.

The full set of models that are equivalent to B would represent isopower alternatives. In many applications this set would be very large, as shown by MacCallum et al. (1993).

Case 3: Any Alternative Model in Which A Is Nested

Given a particular Model A and alternative Model B, isopower alternatives can also be constructed using as the alternative model any model in which A is nested. These alternatives need not have the same number of parameters as B. For example, consider Model B2 in Figure 6. Model B2 has only one more parameter than A, whereas Model B had two more parameters than A. For B2 to serve as an isopower alternative, and holding all of the common parameters between B2 and A to be the same as they were for B, we would need to find a value for the single differentiating parameter, β_1 in Figure 6, such that the resulting power for testing A against B2 would again be .79. A procedure for solving this problem is described in the Appendix. Implementing this procedure in the example here yields the solution $\beta_1 = 0.28$, as shown in Figure 6. For a given alternative model in which A is nested, this approach can be followed to produce isopower alternatives.

Case 4: Any Alternative Model Regardless of Nesting

Case 3 can be extended such that isopower alternatives can be constructed using any alternative Model B regardless of whether A is nested in B. The only constraint would be that B could not be nested in A or equivalent to A (in an equivalent-model sense). For any B that meets these conditions, the procedure described in the Appendix and used under Case 3 in the previous section can be implemented to yield an isopower alternative. For example, Model B3 in Figure 7 has the same number of parameters as A, so there is no nesting relationship (and no model equivalence). This model along with the parameter values shown in Figure 7 would generate an implied covariance matrix Σ_{B3}, and when A is fit to Σ_{B3} the resulting discrepancy function value, $F_{A|B3}$, would match $F_{A|B}$ from our original example, thereby providing an isopower alternative. Again, this procedure can be applied for any alternative model that is not nested in A.
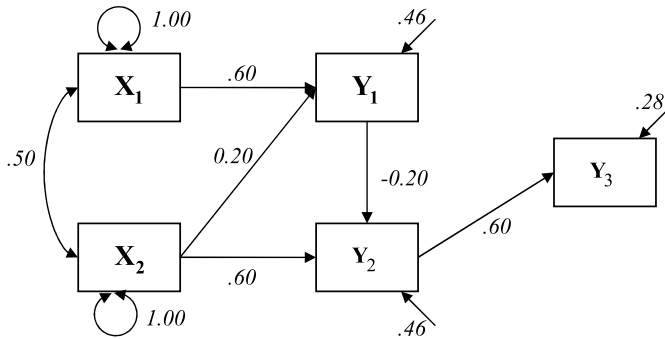
FIGURE 7   Model B3: An isopower alternative in which Model A is not nested.

Case 5: Alternative Covariance Matrices That Are Not Model-Derived

From the framework described earlier it can be seen that the key to defining isopower alternatives is that the same discrepancy function value is obtained when A is fit to the covariance matrices represented by those alternatives. In our running example, any alternative covariance matrix that, when fitted by A, yields a discrepancy function value that matches $F_{A|B} = .126$ would be an isopower alternative. The four previous cases we have described for constructing such alternatives all are model-based, such that different alternative models are constructed and parameter values are determined so as to define a class of isopower alternatives. We consider here the possibility of generating isopower alternatives that are not model-based; that is, constructing isopower alternatives in the form of alternative covariance matrices, all of which would yield the same discrepancy function value, and in turn the same power, when fit by the original Model A.

Surprisingly, such alternatives can be produced easily using a procedure described by Cudeck and Browne (1992) for constructing covariance matrices with a specified lack of model fit. In the present context, given Model A and essentially any assigned parameter values yielding an implied covariance matrix Σ_A, the Cudeck–Browne procedure can be implemented to produce an infinite number of perturbed covariance matrices, $\Sigma_B = \Sigma_A + \Delta$, such that when A is fit to Σ_B, a desired specified discrepancy function value is obtained. (We note that this procedure can break down computationally if the user specifies a discrepancy function value that is too large, representing a very high degree of model misfit. In such situations, though, power analysis would be of little interest because power for detecting extremely poor fit of Model A would virtually always be very high.)

In our running example, Model A with parameter values shown in Figure 8 yields the implied Σ_A shown in Table 1. To generate isopower alternative covariance matrices, we wish to construct perturbed covariance matrices such that when fit by A they will yield a discrepancy function value of .126, as mentioned earlier. The Cudeck–Browne procedure could in theory be used to construct an infinite number of covariance matrices with this property. Table 1 shows two such matrices. Each represents an isopower alternative in the sense that, if either matrix is defined as an alternative hypothesis representing the true state of the population of interest, and if Model A is tested using the likelihood ratio test, the power of that test will be .79. These alter-
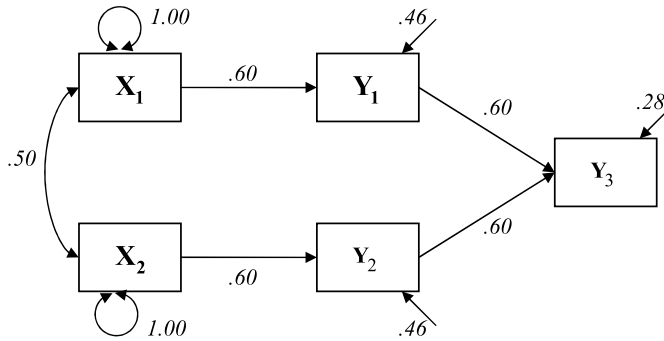
FIGURE 8   Model A with arbitrary specified parameter values.

natives and their difference from Σ_A can be viewed as providing a sense of the degree of general disturbance or perturbation of Σ_A that is associated with a given level of statistical power.

Summary of Isopower Under the Parametric Misspecification Approach

The conventional view of the Satorra–Saris (1985) approach to power analysis is that a Model A is tested under the assumption that a fully specified alternative Model B holds in the population. Under this framework one can determine the power for the test of A, and interpretation is generally stated in terms of the power for detecting particular misspecifications in A, represented by the parameters and their values in B that differentiate B from A. Given

TABLE 1
Implied Covariance Matrix, Σ_A and Two Perturbed Covariance Matrices
Generated by Cudeck and Browne (1992) Procedure

	Variables	y1	y2	y3	x1	x2
Σ_A	y1	0.82				
	y2	0.18	0.82			
	y3	−0.60	−0.60	1.00		
	x1	0.60	0.30	−0.54	1.00	
	x2	0.30	0.60	−0.54	0.50	1.00
$\Sigma_A + \Delta_1$	y1	0.82				
	y2	0.07	0.82			
	y3	−0.53	−0.53	0.92		
	x1	0.60	0.30	−0.50	1.00	
	x2	0.32	0.60	−0.62	0.50	1.00
$\Sigma_A + \Delta_2$	y1	0.82				
	y2	0.28	0.82			
	y3	−0.66	−0.66	1.07		
	x1	0.60	0.41	−0.51	1.00	
	x2	0.39	0.60	−0.60	0.50	1.00

Note. Perturbed covariance matrices were obtained setting the desired minimum discrepancy function value, F at 0.126.

the developments presented in this article, the particular alternative Model B constructed in this sort of power analysis can be recognized to be a single member of an infinitely large class of isopower alternatives. These alternatives need not have a nesting relationship with A, nor even be based on a particular parametric misspecification of A. Recognition of the existence of this class, and possibly exploring some of its particular members, can serve to further illuminate the level and kind of misspecification of A that is associated with a given level of power.

ISOPOWER IN THE MACCALLUM–BROWNE–SUGAWARA FRAMEWORK

Defining H_1 in Terms of Overall Model Fit

Whereas the Satorra–Saris (1985) approach to power analysis is based on defining an alternative to Model A in terms of specific parametric misspecifications of A, MacCallum et al. (1996) instead defined the alternative hypothesis in terms of overall misfit of Model A. Given a Model A to be tested, the null hypothesis of perfect fit can be stated in terms of the population discrepancy function value, F_0, as H_0: $F_{0(A)} = 0$. As noted earlier, under H_0 the test statistic $T = (N - 1)\hat{F}$ will follow a central chi-square distribution with degrees of freedom $d = p(p + 1)/2 - q$.

To define an alternative hypothesis for purposes of power analysis for this test, MacCallum et al. (1996) suggested specifying A to have a particular degree of misfit in the population, rather than to be misspecified by omission of particular parameters that have specified values. To this end MacCallum et al. employed the root mean squared error of approximation (RMSEA) fit index (Browne & Cudeck, 1993; Steiger & Lind, 1980), defined in the population as $\varepsilon = \sqrt{\frac{F_0}{d}}$. From this definition we can see that the null hypothesis of perfect fit of Model A could also be expressed as H_0: $\varepsilon_A = 0$. More generally, we designate the hypothesized value of ε under H_0 as ε_0, so that under the test of perfect fit, $\varepsilon_0 = 0$. Power analysis then requires that one choose an alternative value of RMSEA, designated ε_1, that represents the misfit of Model A under the alternative hypothesis, H_1. Under that alternative we can then compute the noncentrality parameter for the distribution of T under H_1 as $\lambda_1 = (N-1)F_1 = (N-1)d\varepsilon_1^2$. This framework defines the two distributions shown in Figure 3, and power can be computed in the usual way. The result yields the power of the test of the null hypothesis H_0: $\varepsilon_A = 0$ when the alternative hypothesis H_1: $\varepsilon_A = \varepsilon_1$ is in fact true.

For example, suppose we wish to test the null hypothesis of exact fit, H_0: $\varepsilon_A = 0$, for a Model A that has 20 degrees of freedom. And suppose we set $\varepsilon_1 = .05$, implying an alternative hypothesis H_1: $\varepsilon_A = .05$. Letting $N = 200$, the distribution of T under this H_1 would be noncentral chi-square with 20 degrees of freedom and noncentrality parameter $\lambda_1 = (N - 1)d\varepsilon_1^2 = (199)(20)(.05)^2 = 9.95$. Power can then be computed, corresponding to Figure 3, and found to be 0.40. That is, if we test the null hypothesis that RMSEA is zero when the true value of RMSEA is .05, for a model with 20 degrees of freedom, $N = 200$, and $\alpha = .05$, the probability that we will reject the null hypothesis is .40. (MacCallum et al., 1996, provided SAS code for such computations, and Preacher and Coffman, 2006, provide an online calculator.)

MacCallum et al. (1996) generalized this approach to allow for power analysis where the null hypothesis specifies something other than exact fit. For example, Browne and Cudeck (1993) proposed a test of close fit, where the null hypothesis is that H_0: $\varepsilon \leq .05$. One might wish to investigate the power of the test of this H_0 when actual model fit is not close, such as $\varepsilon = .10$. This generalization is straightforward. We simply choose null and alternative values of RMSEA, designated ε_0 and ε_1. We then determine the power for testing H_0: $\varepsilon \leq \varepsilon_0$ when the alternative H_1: $\varepsilon = \varepsilon_1$ is true. Under H_0 the distribution of T is noncentral chi-square with $d = p(p+1)/2 - q$ and noncentrality parameter $\lambda_0 = (N-1)d\varepsilon_0^2$. Under H_1 the distribution of T is noncentral chi-square with the same degrees of freedom and with noncentrality parameter $\lambda_1 = (N-1)d\varepsilon_1^2$. Power is then computed as in Figure 3, with the only difference being that the distribution under H_0 is noncentral rather than central chi-square.

For example, for the test of Model A earlier in this section, if we set $\varepsilon_0 = .05$, $\varepsilon_1 = .10$, and $N = 200$, and we conduct the analysis just described, we find power is .84. This means that if we test the null hypothesis that the RMSEA for Model A is $\leq .05$, when the actual RMSEA is .10, the likelihood that we will reject H_0 is .84. This approach provides estimates of the likelihood of detecting a degree of misfit, regardless of its source, whereas the Satorra–Saris approach focuses on detection of particular model misspecifications.

Isopower Under Overall Misfit of the Tested Model

As just described, for any selected values of ε_0 and ε_1, and holding d, N, and α constant, we can determine the power of the test of H_0: $\varepsilon \leq \varepsilon_0$ when the alternative H_1: $\varepsilon = \varepsilon_1$ is true. Given the outcome of such a power analysis, one way to frame the isopower question in this context would be as follows: Do there exist other pairs of values of ε_0 and ε_1 that would yield the same power? (Note that from this perspective we are allowing both H_0 and H_1 to change to identify isopower alternatives, which represents an important difference from our approach under the Satorra–Saris power analysis framework; we return to this point shortly.)

It is relatively straightforward to identify these isopower alternatives. Given the targeted level of power, and given the choice of any ε_0, and holding d, N, and α fixed, the value of ε_1 can be computed by a method described in the Appendix. To identify a range of isopower alternatives, then, we can choose values of ε_0 across a specified range, and for each selected ε_0 we can compute the corresponding ε_1 that would yield the desired power. Recall that for Model A, with $d = 20$, $N = 200$, and $\alpha = .05$, if we set $\varepsilon_0 = .05$ and $\varepsilon_1 = .10$, then power is .84. The curve in Figure 9 illustrates isopower alternatives for this problem, with each point on the curve representing a different pair of values of ε_0 and ε_1 that yield the same power of .84. For example, if we set $\varepsilon_0 = 0.08$ and $\varepsilon_1 = 0.126$, then power is again .84. Note that we can obtain such a contour for any desired level of power. Figure 10 shows isopower curves for selected power levels of .50, .70, and .90. Such information can be useful in indicating the degree of difference between ε_0 and ε_1 that can be detected with a given probability in testing a model with specified degrees of freedom and using a selected N. For example, the curve for power of .90 shows different pairs of ε_0 and ε_1 that would yield that power; e.g., $\varepsilon_0 = 0.05$ vs. $\varepsilon_1 = 0.105$, or $\varepsilon_0 = 0.10$ vs. $\varepsilon_1 = 0.150$. Based on such information, an investigator might, for example, recognize a need to adjust design features (e.g., sample size) or expectations. Regardless, the mere existence of an infinite number of isopower alternative pairs of ε_0 and ε_1
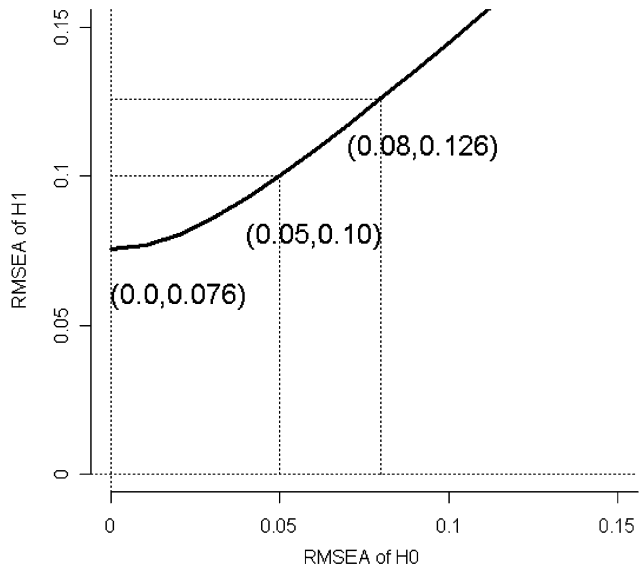
FIGURE 9   Isopower curve showing pairs of root mean squared error of approximation ($\varepsilon_0, \varepsilon_1$) for power of .84, given $N = 200$, $d = 20$, $\alpha = .05$.
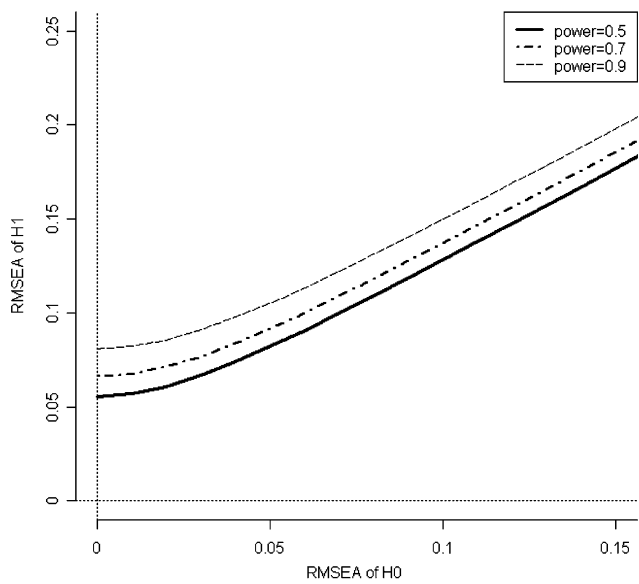
FIGURE 10   Isopower curves showing pairs of root mean squared error of approximation ($\varepsilon_0, \varepsilon_1$) for the same level of power, given $N = 200$, $d = 20$, $\alpha = .05$.

supports the need to interpret any given power calculation as a result associated with only one member of this class.

We noted earlier that the approach just described allows both ε_0 and ε_1 to vary, implying that both H_0 and H_1 vary. In our consideration of isopower alternatives under the Satorra–Saris approach we held H_0 fixed. In the present context of power analysis based on overall model fit, if we hold H_0 fixed (meaning ε_0 is held fixed) there is only one value of ε_1 that yields a given level of power, *unless* we also allow one of the other design factors (d, N, α) to vary. The most interesting approach would be to hold ε_0 fixed and to allow both ε_1 and N to vary. This approach would allow the determination of various combinations of ε_1 and N that would yield the same power as the original test of H_0: $\varepsilon_A \leq \varepsilon_0$ versus the alternative hypothesis H_1: $\varepsilon_A = \varepsilon_1$ where N was specified.

Again recall that in our running example for the test of Model A, if we specify $\varepsilon_0 = .05$, $\varepsilon_1 = .10$, $N = 200$, $d = 20$, and $\alpha = .05$ the resulting power is .84. If we hold fixed ε_0, d, and α, we can identify other combinations of ε_1 and N that yield this same level of power. A method for computing these alternative combinations is described in the Appendix. Application of this method in this example yields the contour for $d = 20$ shown in Figure 11. For instance, holding $\varepsilon_0 = .05$, setting $\varepsilon_1 = 0.083$ and $N = 400$ also yields power of .84. (For illustrative purposes Figure 11 also shows isopower contours for other selected levels of d.) Again, such information serves both to clarify the nature and existence of isopower alternatives and also to assist the investigator in identifying other conditions under which a false H_0 will be rejected at a given level of probability. More general information can be obtained by simply specifying a particular level of power of interest, along with values of ε_0, d, and α, and then computing pairs of values of ε_1 and N that yield the specified power. Table 2 shows an example of
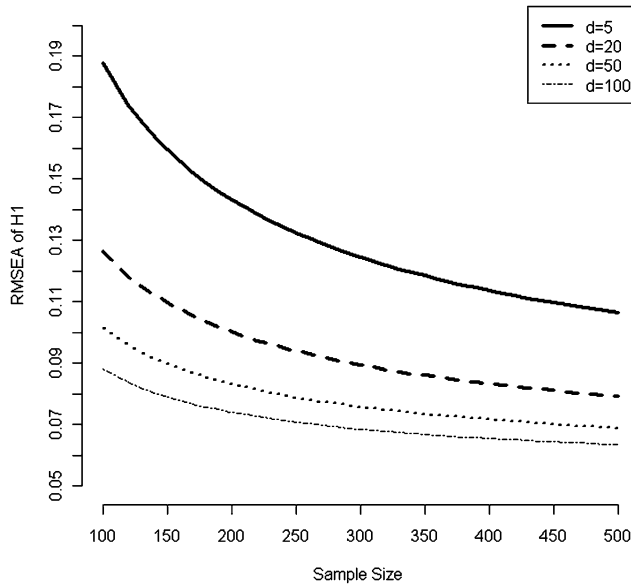
FIGURE 11   Isopower curve showing pairs of ε_1 and N yielding the same level of power $= .84$.

TABLE 2
Pairs of ε_1 and *N* That Produce Power of .84, Holding *df* and α Level Fixed

d	N						
	100	*150*	*200*	*250*	*300*	*400*	*500*
5	.188	.159	.143	.132	.124	.114	.106
10	.153	.131	.118	.110	.104	.096	.091
15	.136	.118	.107	.100	.095	.088	.084
20	.126	.110	.100	.094	.089	.083	.079
25	.119	.104	.095	.090	.086	.080	.076
30	.114	.100	.092	.086	.083	.078	.074
50	.102	.090	.083	.079	.076	.072	.069
75	.093	.083	.077	.074	.071	.068	.066
100	.088	.079	.074	.071	.069	.065	.064

Note. ε_1 was computed setting targeted power at .84, $\alpha = .05$.

such information, illustrating how an investigator can be informed about the level of difference between ε_0 and ε_1 that can be detected at a given N, or the level of N needed to detect various differences between ε_0 and ε_1.

Summary of Isopower Under the Overall Model Fit Approach

Typical use of the MacCallum et al. (1996) approach to power analysis for tests of model fit is based on testing a given Model A with known degrees of freedom. The investigator specifies H_0: $\varepsilon_A \leq \varepsilon_0$ and H_1: $\varepsilon_A = \varepsilon_1$, along with N, d, and α, and computes power as illustrated earlier. The result is then interpreted as the likelihood of rejecting the null hypothesis H_0: $\varepsilon_A \leq \varepsilon_0$ when the alternative hypothesis H_1: $\varepsilon_A = \varepsilon_1$ is actually true. The framework presented earlier shows that the result of such a power analysis in effect selects a single set of conditions from an infinitely large class of conditions that would produce exactly the same outcome. Members of this class can be defined by different combinations of ε_0 and ε_1, holding other elements fixed, or by different combinations of ε_1 and N, again holding other elements fixed. The existence of this class of isopower alternatives impacts interpretation of any particular power analysis, and examination of members of this class can help to inform the investigator about the degree of levels of misfit that are likely to be detected.

DISCUSSION

We have examined the issue of isopower in the context of two different approaches for power analysis in structural equation modeling (SEM). The two approaches use the same underlying statistical distribution theory for determining power (as illustrated in Figure 3), but involve different approaches to defining an alternative hypothesis. The Satorra–Saris (1985) approach defines H_1 in terms of an alternative model assumed to be true in the population, where that alternative model defines particular parametric misspecifications of the model being tested. By contrast, the MacCallum–Browne–Sugawara (1996) approach defines H_1 in terms of a specified

degree of misfit of the model being tested. It is informative to recognize the relationship between the nature of classes of isopower alternatives in the two approaches, as described earlier in this article. Under the Satorra–Saris approach, all members of an isopower class of alternatives yield a constant degree of lack of fit of the tested model in the population. That is, all such isopower alternatives represent alternative models or conditions where, when a tested Model A is fit to the covariance matrix implied by H_1, the same discrepancy function value, $F_{A|B}$, is always obtained. Thus, although the parameter values or actual parameterization of the model defined under H_1 varies, the fit of Model A to the covariance matrix implied under H_1 does not vary. This condition, in turn, then yields the same power for the test of Model A across all such alternatives. Translating this scenario into more familiar lack-of-fit terminology, all such isopower alternatives would yield the same value of common fit measures computed directly from the discrepancy function value, such as the RMSEA index, when fit by Model A.

By contrast, our construction of isopower alternatives under the MacCallum–Browne–Sugawara approach to power analysis allowed for lack of fit of Model A under H_1 to vary, as represented by the alternative value of RMSEA, designated ε_1 earlier in this article. Allowing ε_1 to vary requires that some other factor in the power analysis scenario must also vary for the resulting level of power to be constant. In our study of this problem we showed isopower alternatives obtained by allowing both ε_1 and ε_0 to vary (see Figures 9 and 10), and also by allowing both ε_1 and N to vary (see Figure 11). In summary, our study of the isopower phenomenon under the two general approaches to power analysis in SEM yields two rather different kinds of classes of isopower alternatives. Within the Satorra–Saris approach, all members of such a class represent the same degree of misspecification of a tested Model A, whereas under the MacCallum–Browne–Sugawara approach, members of such a class represent different degrees of lack of fit of Model A coupled with variation in one other quantity in the power analysis computation.

Regardless of which approach is being used, a general principle that applies to all power analysis is relevant here: All power analysis is conditional by nature. The investigator specifies a null hypothesis to be tested, then specifies an alternative hypothesis that is assumed to be true in the population, along with other design and testing features, then computes statistical power for those specified conditions. The result is interpreted as the probability of rejecting H_0 if H_1 is true, and assuming all other specified conditions hold. This article has focused on the fact that different conditions can and will result in the same value for power. In the present context of SEM, we have shown that there can be an infinite number of such combinations of conditions that yield the same power, and also that some computational methods can be applied for identifying many members of this often infinitely large class of alternatives.

In considering implications of this issue for practice, one should recognize the somewhat arbitrary nature of at least some of the conditions and quantities specified in a given power analysis in SEM. The tested model is typically a highly restricted model with development and specification that often involves some uncertainty and exploration. A specified alternative hypothesis also might represent an arbitrary comparison point established for purposes of power analysis. Under the Satorra–Saris approach, such an alternative involves specification of parameter values that are clearly subject to uncertainty. Under the MacCallum–Browne–Sugawara approach the specification of a degree of misfit, represented by ε_1, merely represents an alternative of interest to the researcher rather than a true condition. Rather than consider the set of conditions constructed for a particular power analysis to be rigidly defined and

unchangeable, it is sensible and informative to consider the effect of changing those conditions. This concept raises two possibilities. One is to alter the conditions and examine effects on power, showing how power varies with alterations in various quantities. This idea has been examined earlier in papers including those by Satorra and Saris (1985) and MacCallum et al. (1996). A second approach is to hold power constant and to examine alternative sets of conditions that yield the same power, as investigated in this article. Identification of such isopower alternatives can provide the investigator with further information and insight about the nature and degree of misspecification or misfit that can be detected at a given level of power.

These developments on the issue of isopower have two general implications for investigators using SEM. First, the mere existence of these isopower alternatives should be taken into account when interpreting and reporting the outcome of any power analysis in SEM. More specifically, we suggest that an investigator state that an obtained result for a power computation is not isomorphic with the particular set of conditions specified in the power analysis, and that different sets of conditions likely exist, probably in infinite number, that would yield the same result. We do not view this point as a criticism of power analysis in SEM, but rather as an important aspect of understanding and interpreting a result. From this perspective, the developments in this article can be viewed simply in terms of contributing to the understanding and full interpretation of power analysis in SEM, and not implying a need for further analysis and computation in practice. In fact, it is not our intention to urge routine application of analyses to produce isopower alternatives in empirical applications, but moreso to provide further insight so as to inform interpretation of power analysis in SEM.

Nevertheless, there could well be circumstances where it might be of use to investigators to identify and examine some members of a class of isopower alternatives. So a second implication of this work is that such analyses are feasible and can be informative. As illustrated earlier, for at least some cases this process is not computationally difficult, and the examination of some isopower alternatives might assist the investigator in reaching a richer understanding of power analysis results.[1] As noted earlier, such an examination provides the investigator with a tool for explicitly incorporating into a power analysis the uncertainty associated with specifying conditions for that analysis. For example, in the Satorra–Saris approach, given specification of a tested Model A and alternative Model B as described earlier, the investigator can examine alternative values of the distinguishing parameters representing different degrees of misspecification that could be detected with equal power (Case 1 earlier in this article). Or, still within the context of the Satorra–Saris approach, the investigator could examine the degree of misspecification associated with completely different distinguishing parameters (distinct alternative models) that could be detected with equal power (Cases 2, 3, and 4 earlier). In the MacCallum–Browne–Sugawara approach, given the somewhat arbitrary nature of specification of ε_0 and ε_1, the investigator could examine different degrees of hypothesized and true fit that yield the same level of power (as in Figure 10). Such an extension of a typical power analysis in SEM could more explicitly take into account the conditional and sometimes arbitrary nature of a particular power analysis and provide richer information about the kind and degree of misspecification that can be detected at a given level of power.

---

[1]Software code written to obtain some of the results in this article can be obtained on request from the second author, Taehun Lee, by e-mailing taehun@email.unc.edu.

Thus, although the formal developments in this article were not originally intended for routine computational application in practice, such analyses are feasible and could well be useful to some investigators. Regardless of whether such analyses are conducted in practice, we believe that an awareness and understanding of the phenomenon of isopower can enhance the understanding and interpretation of power analysis results in SEM.

REFERENCES

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Cudeck, R., & Browne, M. W. (1992). Constructing a covariance matrix that yields a specified minimizer and a specified minimum discrepancy function value. *Psychometrika, 57,* 357–369.

Lee, S., & Hershberger, S. (1990). A simple rule for generating equivalent models in covariance structure modeling. *Multivariate Behavioral Research, 25,* 313–334.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130–149.

MacCallum, R. C., Wegener, D. T., Uchino, B. N., & Fabrigar, L. R. (1993). The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin, 114,* 185–199.

Preacher, K. J., & Coffman, D. L. (2006, May). Computing power and minimum sample size for RMSEA [Computer software]. Retrieved December 1, 2009, from http://www.quantpsy.org/

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical recipes in Fortran: The art of scientific computing* (2nd ed.). Cambridge, UK: Cambridge University Press.

Saris, W. E., & Satorra, A. (1993). Power evaluations in structural equation models. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 181–204). Newbury Park, CA: Sage.

Satorra, A., & Saris, W. E. (1985). The power of the likelihood ratio test in covariance structure analysis. *Psychometrika, 50,* 83–90.

Steiger, J. H., & Lind, J. M. (1980, May). *Statistically based tests for the number of common factors.* Paper presented at the annual meeting of the Psychometric Society, Iowa City, IA.

Stelzl, I. (1986). Changing a causal hypothesis without changing the fit: Some rules for generating equivalent path models. *Multivariate Behavioral Research, 21,* 309–331.

APPENDIX

Obtaining Parameter Values for Model B so as to Yield a Specified Power for Test of Model A

Given a Model A nested in a Model B, the problem is to find parameter values for Model B such that a power analysis for a test of fit of Model A (null hypothesis) versus Model B (alternative hypothesis) yields a specified level of power. Equivalently, we wish to find parameter values for Model B such that the test of fit of A when B is true yields a specified value of the discrepancy function, F. Let γ_A be a parameter vector for Model A, and γ_B be a parameter vector for Model B. The covariance matrix implied by Model A is $\Sigma(\gamma_A)$ and the covariance matrix implied by Model B is $\Sigma(\gamma_B)$. When γ_{B0}, a specific value of γ_B, is given, a discrepancy function value, $F_0 = F[\Sigma(\gamma_A), \Sigma(\gamma_{B0})]$ can be obtained by fitting model A to $\Sigma(\gamma_{B0})$. But here our purpose is, given a specified Model A and Model B, to generate $\tilde{\gamma}_B$, a specific value of γ_B that produces a desired discrepancy function value, F^*.

To find such a $\tilde{\boldsymbol{\gamma}}_B$, an iterative procedure is employed. Letting $\boldsymbol{\gamma}_B^0 = [\tilde{\boldsymbol{\gamma}}_{B1}, \boldsymbol{\gamma}_{B2}^0]$ represent an initial value of $\tilde{\boldsymbol{\gamma}}_B = [\tilde{\boldsymbol{\gamma}}_{B1}, \tilde{\boldsymbol{\gamma}}_{B2}]$, the idea is to seek a scaling constant κ_n such that $F[\boldsymbol{\Sigma}(\boldsymbol{\gamma}_A), \boldsymbol{\Sigma}(\tilde{\boldsymbol{\gamma}}_{B1}, \kappa_n \boldsymbol{\gamma}_{B2}^0)] = F^*$ where $F(\cdot, \cdot)$ is the maximum likelihood discrepancy function, and where $\boldsymbol{\gamma}_{B2}$ indicates any one element of parameters of Model B. In the case where Model A is nested within Model B, $\boldsymbol{\gamma}_{B2}$ can be chosen as any one of the parameters additionally specified for Model B. Then $\tilde{\boldsymbol{\gamma}}_{B1}$ is fixed at particular values of parameters in Model B during iteration. Again, in the case where Model A is nested within Model B, $\tilde{\boldsymbol{\gamma}}_{B1}$ contains the same parameters as Model A. The index n represents iteration number. If we consider $y = F[\boldsymbol{\Sigma}(\boldsymbol{\gamma}_A), \boldsymbol{\Sigma}(\tilde{\boldsymbol{\gamma}}_{B1}, \kappa_n \boldsymbol{\gamma}_{B2}^0)] - F^*$ to be a function of κ_n, then this is a straightforward one-dimensional root finding problem. That is, the problem is to find a value of the scaling factor κ such that $y = 0$. In our examples, Brent's method was used to find the appropriate scaling constant. The Brent method is described in detail in section 9.3 of Press, Teukolsky, Vetterling, and Flannery (1992).

It should be noted that $y = F[\boldsymbol{\Sigma}(\boldsymbol{\gamma}_A), \boldsymbol{\Sigma}(\tilde{\boldsymbol{\gamma}}_{B1}, \kappa_n \boldsymbol{\gamma}_{B2}^0)] - F^*$ is a nonlinear equation and therefore, it is possible that there are no roots. In addition, in the case where $\boldsymbol{\gamma}_{B2}$ consists of more than two elements, there exists another source of problem of finding roots. Now $\tilde{\boldsymbol{\gamma}}_{B2}$ is a vector of parameters and $\kappa_n \boldsymbol{\gamma}_{B2}^0$ represents a straight line in the parameter space with direction given by the fixed γ_{B2}^0, and hence the elements of this vector do not change independently. Therefore, possibly some choices of $\boldsymbol{\gamma}_{B2}^0$ will not permit finding a root. In our experience this algorithm has performed very well and has not failed as long as the value of F^* is not set extremely high and the specified parameter values are in a realistic range.

The same procedure can be applied to the situation where Model A and Model B do not have a nesting relationship with each other. Then, the choice of separation of $\boldsymbol{\gamma}_B$ into $\boldsymbol{\gamma}_{B1}$ and $\boldsymbol{\gamma}_{B2}$ can be made arbitrarily. Again, we have found the algorithm to work well in this context as long as Models A and B are not radically different with respect to parametric structure.

Obtaining Pairs of Null and Alternative RMSEA Values to Yield a Specified Level of Power

To compute statistical power in SEM, a total of five pieces of information are needed. That is, the statistical power (π) in SEM is a function of five arguments: significance level (α), sample size (n), degree of freedom (df), noncentrality parameter under the null distribution (λ_0) and noncentrality parameter under the alternative distribution (λ_1), and critical value.

Let χ_c^2 be the critical value under the null hypothesis, given α level, then

$$\pi = f(n, d, \lambda_0, \lambda_1, \chi_c^2)$$

and equivalently,

$$\pi = f(n, d, \varepsilon_0, \varepsilon_1, \chi_c^2)$$

as $\lambda_0 = (n - 1)(d)(\varepsilon_0)^2$ and $\lambda_1 = (n - 1)(d)(\varepsilon_1)^2$.

Therefore, given the target level of power (π), if any four arguments' values are provided, any one remaining argument's value that produces the desired level of power (π) can be

obtained. For example, given the desired power, π, if a researcher provides the information on n, d, ε_0, and χ_c^2, then, by finding the root of the equation

$$f(n, d, \varepsilon_0, \varepsilon_1, \chi_c^2) - \pi = 0$$

the value of ε_1 can be obtained.

More specifically, to get the pairs of ε_0 and ε_1 producing same level of power, say .80, the following steps are needed.

1. Hold d, n, and α fixed and choose a value of ε_0.
2. Compute noncentrality parameter, λ_0, of the chi-square distribution under H_0.
3. Obtain critical value χ_c^2 such that the probability that the chi-square distributed random variable has value greater than χ_c^2 equals α under H_0. That is, $\Pr(\chi^2 > \chi_c^2 | H_0) = \alpha$.
4. Compute the noncentrality parameter λ_1 of the chi-square distribution under H_1, which produces the targeted level of power by solving the equation $\Pr(\chi^2 > \chi_c^2 | H_1) = \pi$.
5. Simple algebraic manipulation of $\lambda_1 = (n-1)(d)(\varepsilon_1)^2$ gives ε_1 value of interest.

With different choice of ε_0 in Step 1, holding d, n, and α fixed, a different value of ε_1 that produces the targeted level of power would be obtained. The pairs of ε_0 and ε_1 computed in this way constitute the isopower alternatives.

Obtaining Pairs of Alternative RMSEA and Sample Size to Yield a Specified Level of Power

The underlying principle of getting pairs of N and ε_1 producing the same power is the same as that of obtaining pairs of ε_0 and ε_1 producing the same level of power as described in the previous section. More specifically,

1. Hold d, ε_0, and α fixed and choose a value of n.
2. Compute noncentrality parameter, λ_0, of the chi-square distribution under H_0.
3. Obtain critical value χ_c^2 such that the probability that the chi-square distributed random variable has value greater than χ_c^2 equals α under H_0. That is, $\Pr(\chi^2 > \chi_c^2 | H_0) = \alpha$.
4. Compute the noncentrality parameter λ_1 of the chi-square distribution under H_1, which produces the targeted level of power by solving the equation $\Pr(\chi^2 > \chi_c^2 | H_1) = \pi$.
5. Simple algebraic manipulation of $\lambda_1 = (n-1)(d)(\varepsilon_1)^2$ gives ε_1 value of interest.

With a different choice of n in Step 1, holding d, ε_0, and α fixed, a different value of ε_1 that produces the targeted level of power would be obtained. The pairs of n and ε_1 computed in this way constitute isopower alternatives.