of model, I present a sample application including the source code, printout, and results section. Chapter 8 presents some "tricks of the trade" for structural equation modeling, including the use of single indicator latent variables and reducing the cognitive complexity of models.

Although a comprehensive understanding of structural equation modeling is a worthwhile goal, I have focused in this book on the three most common forms of analysis. In doing so, I have "glossed over" many of the refinements and types of analyses that can be performed within a structural equation modeling framework. I have also tried to stay away from features of LISREL VIII (Jöreskog & Sörbom, 1992) that are implementation-dependent. For example, I do not discuss the implementation of the SIMPLIS language or the graphical interface available in LISREL VIII. Although this choice may limit the current presentation with respect to LISREL VIII, it also makes the book relevant to users of older versions of LISREL.

When all is said and done, the intent of this book is to give a "user-friendly" introduction to structural equation modeling. The presentation is oriented to researchers who want or need to use structural equation modeling techniques to answer substantive research questions. Those interested in a more mathematical presentation are referred to the ever growing body of literature on the derivation and implementation of structural equation models.

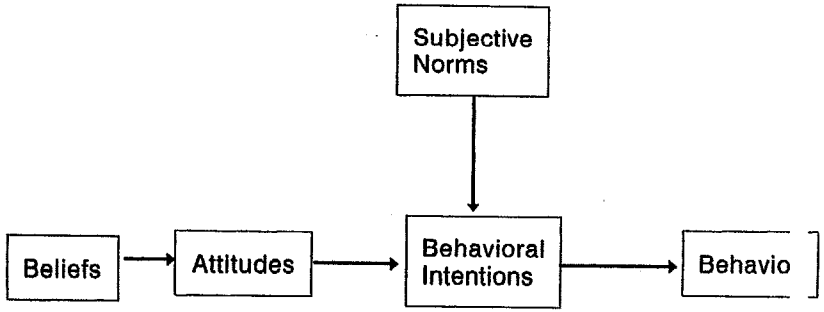
CHAPTER 2 *Structural Equation Models*

Theory and Development

To begin, let us consider what we mean by the term *theory*. At one level, a theory can be thought of as an explanation of why variables are correlated (or not correlated). Of course, most theories in the social sciences go far beyond the description of correlations to include hypotheses about causal relations. A necessary but insufficient condition for the validity of a theory would be that the relationships (i.e., correlations/covariances) among variables are consistent with the propositions of the theory.

For example, consider Fishbein and Ajzen's (1975) well-known theory of reasoned action. In the theory (see Figure 2.1), the best predictor of behavior is posited as being the intention to perform the behavior. In turn, the intention to perform the behavior is thought to be caused by (a) the individual's attitude toward performing the behavior and (b) the individual's subjective norms about the behavior. Finally, attitudes toward the behavior are thought to be a function of the individual's beliefs about the behavior. This simple presentation of the theory is sufficient to generate some expectations about the pattern of correlations between the variables referenced in the theory.

If the theory is correct, then one would expect that the correlation between behavioral intentions and behavior and the correlation between beliefs and attitudes should be stronger than the correlations between



attitudes and behavior or subjective norms and behavior. Correspor l-
ingly, the correlations between beliefs and behavioral intentions a d
beliefs and behavior should be the weakest correlations. With referer e
to Figure 2.1, the general principle is that if the theory is correct, th n
direct and proximal relationships should be stronger than more dis il
relationships.

As a simple test of the theory, one could collect data on behavi r,
behavioral intentions, attitudes, subjective norms, and beliefs. If t e
theory is correct, one would expect to see the pattern of correlatio s
described above. If the actual correlations do not conform to the patte i,
one could reasonably conclude that the theory was incorrect (i.e., t e
model of reasoned action did not account for the observed correlation .

Note that the converse is not true. Finding the expected pattern f
correlations would not imply that the theory is right, only that it s
plausible. There might be other theories that would result in the sar e
pattern of correlations (e.g., one could hypothesize that behavior caus s
behavioral intentions, which in turn cause attitudes and subjective norm .
As noted earlier, finding the expected pattern of correlations is a
necessary but not sufficient condition for the validity of the theory.

Although the above example was a simple one, it illustrates the log c
of structural equation modeling. In essence, structural equation mod -
ing is based on the observations that (a) every theory implies a set f
correlations and (b) if the theory is valid, then the theory should be ab e
to explain or reproduce the patterns of correlations found in tl e
empirical data.

The Process of Structural Equation Modeling

The remainder of this chapter is organized according to a linear "model" of structural equation modeling. Although linear models of the research process are notoriously suspect (McGrath, Martin, & Kukla, 1982) and may not reflect actual practice, the heuristic has the advantage of drawing attention to the major concerns, issues, and decisions involved in developing and evaluating structural equation modeling. Bollen and Long (1993, pp. 1-2) describe the five stages characteristic of most applications of structural equation modeling:

1. model specification,
2. identification,
3. estimation,
4. testing fit, and
5. respecification.

For presentation purposes, I will defer much of the discussion of testing fit until the next chapter.

Model Specification

Structural equation modeling is inherently a confirmatory technique. That is, for reasons that will become clear as the discussion progresses, the methods of structural equation modeling are ill suited for the exploratory identification of relationships. Rather, the foremost requirement for any form of structural equation modeling is the a priori specification of a model. The propositions composing the model are most frequently drawn from previous research or theory (Bollen & Long, 1993), although the role of informed judgment, hunches, and dogmatic statements of belief should not be discounted. However derived, the purpose of the model is to explain why variables are correlated in a particular fashion. Bollen (1989, p. 1), for example, presents the fundamental hypothesis for structural equation modeling as:

$$\Sigma = \Sigma(\Theta)$$

where Σ is the observed population covariance matrix, Θ is a vector of model parameters, and $\Sigma(\Theta)$ is the covariance matrix implied by the model. When the equality expressed in the equation holds, the model is said to "fit" the data. Thus, the goal of structural equation modeling is to explain the patterns of covariance observed among the study variables.

In essence, then, a model is an explanation of why two (or more) variables are related (or not). In undergraduate statistics courses, we often harp on the observation that a correlation between *X* and *Y* has at least three possible interpretations (i.e., *X* causes *Y*, *Y* causes *X*, or *X* and *Y* are both caused by a third variable *Z*). In formulating a model, you are choosing one of these explanations, in full recognition of the fact that either of the remaining two might be just as good, or better, explanations.

It follows from these observations that the "model" used to explain the data cannot be derived from that data. For any covariance or correlation matrix, one can always derive a model that provides a perfect fit to the data. Rather, the power of structural equation modeling derives from the attempt to assess the fit of theoretically derived predictions to the data.

It might help at this point to consider two types of variables. In any study, we have variables we want to explain or predict. We also have variables that we think will offer the explanation or prediction we desire. The former are known as *endogenous* variables, whereas the latter are *exogenous* variables. Exogenous variables are considered to be the starting points of the model. We are not interested in how the exogenous variables came about. Endogenous variables may serve as both predictors and criteria, being predicted by exogenous variables and predicting other endogenous variables. A model, then, is a set of theoretical propositions that link the exogenous variables to the endogenous variables and the endogenous variables to one another. Taken as a whole the model explains both what relationships we expect to see in the data and what relationships we do not expect to emerge.

It is worth repeating that the fit of a model to data, in itself, conveys no information about the validity of the underlying theory. Thankfully the misnomer (Breckler, 1990) "causal modeling" appears to have passed out of fashion, with the recognition that structural equation models do not assess or "prove" causality any more than the application of any statistical technique conveys information about the causal relations in the data (Williams, 1995).

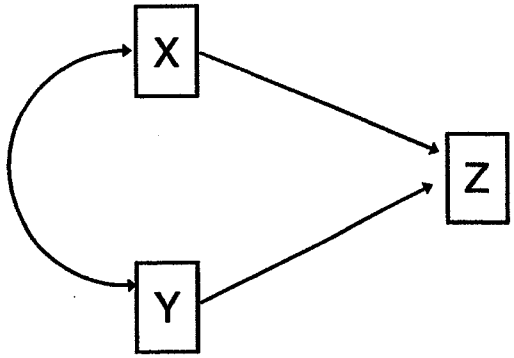


Figure 2.2.

Although the hypotheses underlying model development may be causal in nature, assessing the fit of a model does not provide a basis for causal inference (Brannick, 1995; Kelloway, 1995; Williams, 1995). The conditions necessary for causal inference in structural equation modeling are presented by James and colleagues (James, Mulaik, & Brett, 1982) and are more briefly summarized by Bollen (1989) as three main conditions: (a) association, (b) isolation (the inclusion of all relevant predictors), and (c) the establishment of causal direction. Meeting these conditions for causal inference is more a matter of study design than of statistical technique.

Path diagrams. Most frequently, the structural relations that form the model are depicted in a path diagram in which variables are linked by unidirectional arrows (representing causal relations) or bidirectional curved arrows (representing noncausal, or correlational, relationships).[1] Consider three variables *X*, *Y*, and *Z*. A possible path diagram depicting the relationships among the three is given in Figure 2.2.

The diagram presents two exogenous variables (*X* and *Y*) that are assumed to be correlated (curved arrow). Both variables are presumed to cause *Z* (unidirectional arrows).

Now consider adding a fourth variable, *Q*, with the hypotheses that *Q* is caused by both *X* and *Z*, with no direct effect of *Y* on *Q*. The path diagram representing these hypotheses is presented in Figure 2.3.

Three important assumptions underlie path diagrams. First, it is assumed that all of the proposed causal relations are linear. Although

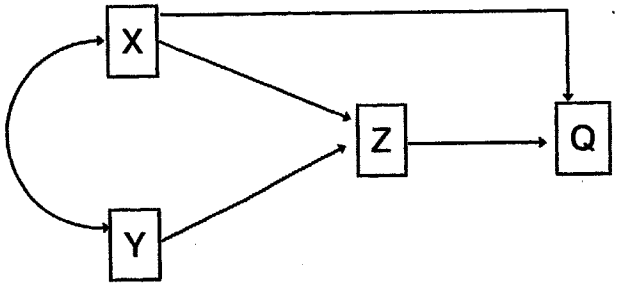


Figure 2.3.

there are ways of approximating nonlinear relations in structural equation modeling (see Kenny & Judd, 1984), for the most part we are concerned only with linear relations. Second, path diagrams are assumed to represent all the causal relations between the variables. It is just as important to specify the causal relationships that do exist as it is to specify the relationships that do not. Finally, path diagrams are based on the assumption of causal closure; this is the assumption that all causes of the variables in the model are represented in the model. That is, any variable thought to cause two or more variables in the model should in itself be part of the model. Failure to actualize this assumption results in misleading and often inflated results (which economists refer to as specification error). In general, we are striving for the most parsimonious diagram that (a) fully explains why variables are correlated and (b) can be justified on theoretical grounds.

Finally, it should be noted that one can also think of factor analysis as a path diagram. The common factor model on which all factor analyses are based states that the responses to an individual item are a function of (a) the trait that the item is measuring and (b) error. Another way to phrase this is that the observed variables (items) are a function of both common factors and unique factors.

For example, consider the case of six items that are thought to load on two factors (which are oblique). Diagrammatically, we can represent this model as shown in Figure 2.4. Note that this is the conceptual model that we have when planning a factor analysis. As will be explained in greater detail later, the model represents the confirmatory factor analysis model and not the model commonly used for exploratory factor analysis.

In the diagram, F1 and F2 are the two common factors. They are also referred to as *latent variables* or unobserved variables because they are

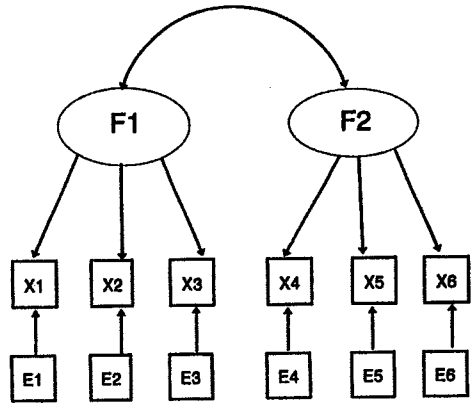


Figure 2.4.

not measured directly. Note that it is common to represent latent variables in ovals or circles. X1 . . . X6 are the *observed* or *manifest variables* (test items, sometimes called indicators), whereas E1 . . . E6 are the residuals (sometimes called unique factors or error variances). Thus, although most of this presentation focuses on path diagrams, all the material is equally relevant to factor analysis, which can be thought of as a special form of path analysis.

Converting the path diagram to structural equations. Path diagrams are most useful in depicting the hypothesized relations because there is a set of rules that allow one to translate the diagram into a series of structural equations. The rules, initially developed by Sewall Wright (1934), allow one to write a set of equations that completely define the observed correlations matrix.

The logic and rules for path analysis are quite straightforward. The set of arrows constituting the path diagram include both simple and compound paths. A *simple path* (e.g., $X \rightarrow Y$) represents the direct relationship between two variables (i.e., the regression of Y on X). A compound path (e.g., $X \rightarrow Y \rightarrow Z$) consists of two or more simple paths. The value of a compound path is the product of all the simple paths constituting the compound path. Finally, and most important for our purposes, the correlation between any two variables is the sum of the simple and compound paths linking the two variables.

Given this background, Sewall Wright's rules for decomposing correlations are these:

1. After going forward on an arrow, the path cannot go backward. The path can, however, go backward as many times as necessary prior to going forward.
2. The path cannot go through the same construct more than once.
3. The path can include only one curved arrow.

Consider, for example, three variables, *A*, *B*, and *C*. Following psychological precedent, I measure these variables in a sample of 100 undergraduates and produce the following correlation matrix:

	A	B	C
A	1.00		
B	.50	1.00	
C	.65	.70	1.00

I believe that both *A* and *B* are causal influences on C. Diagrammatically, my model might look like the model shown in Figure 2.5.

Following the standard rules for computing path coefficients, I can write a series of structural equations to represent these relationships. By solving for the variables in the structural equations, I am computing the path coefficients (the values of the simple paths).

$$c = .5$$

$$a + cb = .65 \qquad (2.1)$$

$$b + ca = .70 \qquad (2.2)$$

Note that three equations completely define the correlation matrix. That is, each correlation is thought to result from the relationships specified in the model. Those who still recall high school algebra will recognize that I have three equations to solve for three unknowns; therefore, the solution is straightforward. Because I know the value of *c* (from the correlation matrix), I begin by substituting *c* into Equations 2.1 and 2.2. Equation 2.1 then becomes

$$a + .5b = .65 \qquad (2.1.1)$$

and Equation 2.2 becomes

$$b + .5a = .70 \qquad (2.2.1)$$

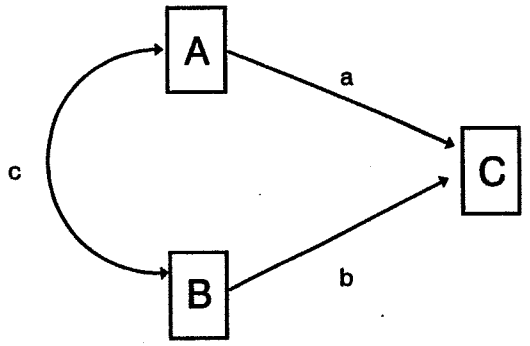


Figure 2.5.

To solve the equations, one can multiply Equation 2.2.1 by 2 (resulting in Equation 2.2.2) and then subtract Equation 2.1.1 from the result.

$$2b + a = 1.4 \qquad (2.2.2)$$

$$-(.5b + a = .65 \qquad (2.1.1)$$

$$= 1.5b \quad = .75 \qquad (2.3)$$

From Equation 2.3, we can solve for *b*: $b = .75/1.5 = .50$. Substituting *b* into either Equation 2.2.1 or Equation 2.1.1 results in $a = .40$. Thus, the three path values are $a = .40$, $b = .50$, and $c = .50$.

These numbers are standardized partial regression coefficients or beta weights and are interpreted exactly the same as betas derived from multiple regression analyses. Indeed, a simpler method to derive the path coefficients *a* and *b* would have been to use a statistical package to conduct an ordinary least squares regression of *C* on *A* and *B*. The important point is that any model implies a set of structural relations among the variables. These structural relations can be represented as a set of structural equations and, in turn, imply a correlation (or covariance) matrix.

Thus, a simple check on the accuracy of the solution is to work backwards. Using the estimates of structural parameters, we can calculate the correlation matrix. If the matrix is the same as the one we started out with, we have reached the correct solution. Thus,

$$c = .5$$

$$a + cb = .65$$

$$b + ca = .70$$

and we have calculated that $b = .5$ and $a = .4$. Substituting into the second equation above, we get $.4 + .5 \times .5 = .65$ or $.4 + .25 = .65$. For the second equation, we get $.5 + .5 \times .4 = .70$, or $.5 + .20 = .70$. In this case, our model was able to reproduce the correlation matrix. That is, we were able to find a set of regression or path weights for the model that can replicate the original, observed, correlations.

Identification

As illustrated by the foregoing example, application of structural equation modeling techniques involves the estimation of unknown parameters (e.g., factor loadings or path coefficients) based on observed covariances/correlations. In general, issues of identification deal with whether a unique solution for the model (or its component parameters) can be obtained (Bollen, 1989). Models and/or parameters may be underidentified, just-identified, or overidentified (Pedhazur, 1982).

In the example given above, the number of structural equations composing the model exactly equals the number of unknowns (i.e., three unknowns and three equations). In such a case, the model is said to be *just-identified* (because there is just one correct answer). A just-identified model will always provide a unique solution (i.e., set of path values) that will be able to perfectly reproduce the correlation matrix. A just-identified model is also referred to as a saturated model (Medsker et al., 1994).

A necessary, but insufficient, condition for the identification of a structural equation model is that one cannot estimate more parameters than there are unique elements in the covariance matrix. Bollen (1989) refers to this as the "t rule" for model identification. Given a $k \times k$ covariance matrix (where k is the number of variables), there are $k \times (k-1)/2$ unique elements in the covariance matrix. Attempts to estimate exactly $k \times (k-1)/2$ parameters results in the just-identified or "saturated" (Medsker et al., 1994) model. Only one unique solution is obtainable for the just-identified model, and the model always provides a perfect fit to the data.

When the number of unknowns exceeds the number of equations, the model is said to be *underidentified*. This is a problem because the model parameters cannot uniquely ascertained; there is no unique solution. Consider, for example, the solution to the equation $X + Y = 10$. There are no two unique values for X and Y that solve this equation (there are, however, an infinite number of possibilities).

Last, and most important, when the number of equations exceeds the number of unknowns, the model is *overidentified*. When models are overidentified, there are a number of unique solutions, and the task in most applications of structural equation modeling techniques is to find the solution that provides the best fit to the data. Thus, the identification of a structural equation model is purely a matter of the number of estimated parameters (Bollen, 1989).

The ideal situation for the social scientist is to have an overidentified model. If the model is underidentified, no solution is possible. If the model is just-identified, then there is one set of values that completely fit the observed correlation matrix. That matrix, however, also contains many sources of error (e.g., sampling error, measurement error). In an overidentified model, there are a number of possible solutions, and the task is to select the one that comes closest to explaining the observed data within some margin of error. We always, therefore, want our models to be overidentified.

Although it is always possible to "prove" that your proposed model is overidentified (see Long, 1983a, 1983b for examples), the procedures are cumbersome and involve extensive calculations. Overidentification of a structural equation model is achieved by placing two types of restrictions on the model parameters to be estimated.

First, researchers assign a direction to parameters. In effect, positing a model based on one-way causal flow restricts half of the posited parameters to be zero. Models incorporating such a one-way causal flow are known as *recursive models*. Bollen (1989) points out that recursiveness is a sufficient condition for model identification. That is, as long as all the arrows are going in the same direction, the model is identified. Moreover, in the original formulation of path analysis, where path coefficients are estimated through OLS regression (Pedhazur, 1982), recursiveness is a required property of models. Recursive models, however, are not a necessary condition for identification, and it is possible to estimate identified nonrecursive models (i.e., models that incorporate reciprocal causation) using programs such as LISREL.

Second, researchers achieve overidentification by setting some parameters to be fixed to a predetermined value. Typically, values of specific

parameters are set to zero. Earlier, in the discussion of model specification, I made the point that it is important for researchers to consider (a) what paths will be in the model and (b) which paths are not in the model. By "not in the model," I am referring to the setting of certain paths to zero. For example, in the theory of reasoned action presented earlier (see Figure 2.1), several potential paths (i.e., from attitudes to behavior, from norms to behavior, from beliefs to intentions, from beliefs to norms, and from beliefs to behavior) were set to zero to achieve overidentification. Had these paths been included in the model, the model would be just-identified.

Estimation and Fit

If the model is overidentified, then, by definition, there are an infinite number of solutions. Moreover, given moderately complex models, solving the structural equations by hand would quickly become a formidable problem. Indeed, the growing popularity of structural equation modeling is probably most attributable to the availability of software packages such as LISREL that are designed to solve sets of structural equations.

LISREL solves these equations (as do most similar programs) by using numerical methods to estimate parameters. In particular, LISREL solves for model parameters by a process of iterative estimation. To illustrate the process of iterative estimation, consider a common children's guessing game.

When I was a boy, we played a game called hot, hotter, hottest. In one version of the game, one child would pick a number and the other child would attempt to guess the number. If the guess was close, the guesser was "getting hotter." If the guess was way off, the guesser was "getting colder." By a simple process of informed trial and error, you could almost always guess the number.

This is precisely the process LISREL uses to estimate model parameters. The program starts by taking a "guess" at the parameter values. It then calculates the implied covariance matrix (the covariance matrix that would result from that set of model parameters). The implied covariance matrix is then compared to the observed covariance matrix (i.e., the actual data) to see how "hot" the first guess was. If the guess was right (i.e., if the implied and actual covariance matrices are very similar), the process stops. If the guess was wrong, then LISREL adjusts the first guess (the starting values) and checks again. This process of iterative estimation continues until some fitting criterion has been achieved (the solution is "red hot").

How does LISREL know when it is "red hot"—that is, when the correct answer is obtained? In general, the user specifies a fitting criterion (a mathematical function) that the program tries to minimize. When repeated iterations fail to minimize the fitting criterion, LISREL grinds to a halt and reports the last solution it estimated. Three very common fitting criteria are ordinary least squares (OLS), generalized least squares (GLS), and maximum likelihood (ML). In matrix notation, these criteria are defined as

$$\text{OLS} = \text{tr}(S - C)^2$$

$$\text{GLS} = 1/2\ \text{tr}[(S - C)S^{-1}]^2$$

$$\text{ML} = \ln|C| - \ln|S| + \text{tr}\ SC^{-1} - m$$

where tr = trace (sum of the diagonal elements),
S = covariance matrix implied by the model,
C = actual covariance matrix,
ln = natural logarithm, and
| | indicates the determinant (index of generalized variance) of a matrix.

Although the specifics of the equations are not important, it is important to note that each criterion attempts to minimize the differences between the implied and observed covariance matrices. That is, each equation has as its basis the direct comparison of S and C. When the observed and predicted covariance matrices are exactly the same, all the above criteria will equal 0. Conversely, when the matrices are different, the value of the fitting function gets larger. Thus, the goal of the iterative estimation procedure used by LISREL is to minimize the fitting function specified by the user.

Because of the complexity of the subject, we will defer further discussion of assessing model fit until the next chapter. Three additional issues regarding model estimation should be noted, however: the choice of estimators, the choice of data type, and sample size requirements.

Choice of Estimators

Perhaps the most widely used type of estimation is maximum likelihood, followed by generalized least squares (Anderson & Gerbing, 1988; Kelloway, 1996). Maximum likelihood estimation is so popular

that researchers seem to equate using LISREL with doing maximum likelihood estimation (Kelloway, 1996), even though LISREL allows a variety of estimation techniques.

Maximum likelihood estimators are known to be consistent and asymptotically efficient in large samples (Bollen, 1989). The popularity of these methods, however, is more likely attributable to the fact that (under certain conditions) the minimum of the fitting criterion multiplied by $N - 1$ (where N is the number of observations) is distributed as χ^2. For maximum likelihood estimation, if we have a large sample and are willing to assume (or show) that the observed variables are multivariate normal, then the χ^2 test is reasonable. If we have a large sample but are not willing to assume multivariate normality, then GLS estimation is the method of choice.

Although most readers will be familiar with Ordinary Least Squares estimation as used in multiple regression, maximum likelihood and generalized least squares estimation are slightly different. Ordinary least squares is known as a partial information technique, whereas both maximum likelihood estimation and generalized least squares are full information techniques. To understand this distinction, consider the diagram of the Fishbein and Ajzen (1975) theory of reasoned action presented in Figure 2.1.

As mentioned earlier, one way to solve for the path values in this model would be to conduct a series of three regression analyses. That is, one would regress the following:

1. behavior on behavioral intentions,
2. behavioral intentions on subjective norms and attitudes, and
3. attitudes on beliefs.

If you were to follow this strategy, then each path value would be estimated independently of the others. If, for example, you made a mistake in calculating the value for the path relating beliefs to attitudes, that would not affect the value of the path leading from behavioral intentions to behavior.

In contrast, maximum likelihood estimation is a full information technique. Simply put, in using maximum likelihood estimation, one estimates all the parameters (i.e., path values) simultaneously. In this case, an error in one value (e.g., a poorly specified part of the model) will be reflected in every parameter estimated.

The consequences of using full information techniques as opposed to partial information techniques are potentially serious (Brannick, 1995) but as yet poorly understood (Williams, 1995). In particular, the advantages of full information techniques (i.e., the availability of hypothesis testing) may outweigh the potential disadvantages. Moreover, under many circumstances, maximum likelihood estimation and ordinary least squares will result in identical estimates, lending weight to the use of full information techniques.

Choice of Data

Up until this point, I have been using the terms *correlation matrix* and *covariance matrix* interchangeably. The two matrices are very similar. A covariance matrix has measures of covariance in the off-diagonal positions, with measures of variance in the main diagonal. A correlation matrix is simply a standardized covariance matrix (i.e., because all the variables are standardized, a correlation matrix has 1s in the main diagonal).

Although the matrices are very similar, the standardization of variables in constructing a correlation matrix removes important information about the scale of measurement of individual variables from the data. This is an important concern because many types of structural equation models are not scale invariant (Cudeck, 1989). That is, one will get different results depending on whether the analysis is based on a correlation or a covariance matrix.

Some authors have suggested that the choice of a correlation or a covariance matrix for input is based on both theoretical concerns and the preferences of some disciplines. Theoretically, if one is concerned only with the pattern of relationships among variables, a correlation matrix is an appropriate choice. Indeed, use of a correlation matrix simplifies interpretation of the results by rescaling all variables to have unit variance. Moreover, use of the correlation matrix may result in more conservative estimates of parameter significance (generally held to be desirable in statistics).

Having said this, use of the covariance matrix is strongly recommended in virtually all instances. Structural equation models are not always scale free—thus, a model that fits the correlation matrix may not fit the covariance matrix. Moreover, the hypothesis tests available in structural equation modeling are based on the assumption that one is

analyzing a covariance matrix. Thus, although there may be valid reasons to analyze a correlation matrix, use of the covariance matrix generally is recommended. This general recommendation applies with equal force to all the types of analysis discussed here.

Sample Size

Although it may not have been obvious up until this point, structural equation modeling is very much a large sample technique. Both the estimation methods (e.g., maximum likelihood) and tests of model fit (e.g., the χ^2 test) are based on the assumption of large samples. Several authors have presented guidelines on the definition of "large" (e.g., Anderson & Gerbing, 1984; Bentler & Chou, 1987; Marsh, Balla, & MacDonald, 1988).

In general, it seems that a sample size of at least 200 observations would be an appropriate minimum. For example, it is commonly recommended that models incorporating latent variables require at least a sample size of 100 observations, although parameter estimates may be inaccurate in samples of less than 200 (Marsh et al., 1988). Boomsma (1983) recommends a sample size of approximately 200 for models of moderate complexity. Taking a somewhat different approach, Bentler and Chou (1987) have suggested that the ratio of sample size to estimated parameters be between 5:1 and 10:1 (similar to frequently cited guidelines for regression analyses, e.g., Tabachnick & Fidell, 1996).

Model Modification

Perhaps no aspect of structural equation modeling techniques is more controversial than the role of model respecification. The goal of model respecification is to improve either the parsimony or the fit of the model (MacCallum, 1986). Thus, respecification typically consists of one of two forms of model modification. First, researchers may delete nonsignificant paths from their models in a "theory-trimming" (Pedhazur, 1982) approach. Second, researchers may add paths to the model based on the empirical results.

Although model respecification frequently is included in descriptions of the modeling process (e.g., Bollen & Long, 1993), there are several problems with specification searches. Perhaps most important, the available data suggest that specification searches typically do not retrieve the actual model (MacCallum, 1986). Moreover, because specification

searches are conducted post hoc and are empirically rather than theoretically derived, model modifications based on such searches must be validated on an independent sample. As James and James (1989) point out, it is perfectly acceptable to modify the model and assess the fit of the model based on data from one sample; it is the interpretation of such model modifications that is suspect. When models are modified and reassessed on the same data, parameters added to or deleted from the model cannot be said to be confirmed.

Aside from the exploratory nature of model respecifications, there is considerable doubt about the meaning of parameters added to a model based on a specification search. Certainly, there are examples in the literature (and in my own work; see Barling, Kelloway, & Bremmerman, 1991) of adding substantively uninterpretable parameters (e.g., covariances among error terms) to a model to improve the fit of the model. Such parameters have been termed "wastebasket" parameters (Browne, 1982), and there is little justification for their inclusion in structural models (Kelloway, 1995, 1996).

It is tempting to conclude, as I have previously (Kelloway, 1996), that parameters that can be assigned a substantive meaning are "legitimate" additions to a structural model during a specification search. Steiger (1990, p. 175) pointed to the flaw in this conclusion when he questioned, " What percentage of researchers would find themselves unable to think up a 'theoretical justification' for freeing a parameter? In the absence of empirical information to the contrary, I assume that the answer . . . is 'near zero.' "

Although replication of model modifications on an independent sample is commonly recognized to be an appropriate strategy, it should be noted that there are also problems with this strategy. Perhaps most important, because the empirically driven respecification of model parameters capitalizes on chance variations in the data, the results of such replication efforts may be inconsistent (MacCallum, Roznowski, & Necowitz, 1992). Thus, there are both conceptual and empirical problems with the practice of respecifying models and, at best, such respecifications provide limited information.

So what do you do if your model doesn't fit the data? One solution to an ill-fitting model is to simply stop testing and declare the theory that guided model development to be wrong. This approach has the advantage of conforming to a classical decision-making view of hypothesis testing; that is, you have a hypothesis, you perform a test, and you either accept or reject the hypothesis. The disadvantage of this approach is, of course, that one does not gain any insight into what the "correct"

(or at least one plausible) theory might be. In particular, there i information in the data you have collected that you may not be using t its fullest advantage.

A second approach to an ill-fitting model is to use the availabl information to try to generate a more appropriate model. This is th "art" of model modification—changing the original model to fit th data. Although model modification is fraught with perils, I do no believe that anyone has ever "gotten it right" on the first attempt a model fitting. Thus, the art of model fitting is to understand the danger and try to account for them when you alter your model based o empirical observations.

The principal danger in post hoc model modification is that thi procedure is exploratory and involves considerable capitalization o chance. Thus, you might add a path to a model to make it fit the dat only to find that you have capitalized on chance variation within you sample and the results will never be replicated in another sample. Ther are at least two strategies for minimizing this problem.

First, try to make model modifications that have some semblance o theoretical consistency (bearing in mind Steiger's comments about ou ability to rationalize). If there are 20 studies suggesting that job satisfac tion and job performance are unrelated, do not hypothesize a patl between satisfaction and performance just to make your model fit Second, as with any scientific endeavor, models are worthwhile onl when they can be replicated in another sample. Post hoc modification to a model should always be (a) identified as such and (b) replicated i another sample.[2]

Notes

1. It also helps to remember that in path diagrams, the hypothesized causal "flow is traditionally from left to right (or top to bottom); that is, the independent (exogenous variables or predictors are on the left (top), and the dependent (endogenous) variable or criteria are on the right (bottom).

2. Note that the use of a holdout sample is often recommended for this purpose. Se aside 25% of the original sample, then test and modify the model on the remaining 75% When you have a model that fits the data on the original 75%, test the model on th remaining 25%. Although this procedure does not always result in replicated finding it can help identify which paths are robust and which are not.

CHAPTER 3 *Assessing Model Fit*

Perhaps more has been written about the assessment of model fit than any other aspect of structural equation modeling. Indeed, many researchers are attracted to structural equation modeling techniques because of the availability of global measures of model fit (Brannick, 1995). In practice, such measures often are used as an omnibus test of the model whereby one first assesses global fit before proceeding to a consideration of the individual parameters composing the model (Jöreskog, 1993). A variety of fit indices are currently available to researchers wishing to assess the fit of their models, and it is instructive to consider exactly what we mean when we claim that a model "fits" the data.

At least two traditions in the assessment of model fit are apparent (Tanaka, 1993): the assessment of the *absolute* fit of the model and the assessment of the *comparative* fit of the model. The assessment of the comparative fit of the model may be further subdivided into the assessment of comparative fit and *parsimonious* fit. The assessment of absolute fit is concerned with the ability of the model to reproduce the actual covariance matrix. The assessment of comparative fit is concerned with comparing two or more competing models to assess which provides the better fit to the data.

The assessment of parsimonious fit is based on the recognition that one can always obtain a better fitting model by estimating more parameters. (At the extreme, one can always obtain a perfect fit to the data by estimating the just-identified model containing all possible parameters.) Thus, the assessment of parsimonious fit is based on the idea of a