5.2 Graphs and Model Testing

In 1919, Wright developed his "method of path coefficients," which allows researchers to compute the magnitudes of cause-effect relationships from correlation measurements provided the path diagram represents correctly the causal processes underlying the data. Wright's method consists of writing a set of equations, one for each pair of variables (X_i, X_j), and equating the (standardized) correlation coefficient ρ_{ij} with a sum of products of path coefficients and residual correlations along the various paths connecting X_i and X_j. One can then attempt to solve these equations for the path coefficients in terms of the observed correlations. Whenever the resulting equations give a unique solution to some path coefficient p_{mn} that is independent of the (unobserved) residual correlations, that coefficient is said to be *identifiable*. If every set of correlation coefficients ρ_{ij} is compatible with some choice of path coefficients then the model is said to be *untestable* or *unfalsifiable* (also called *saturated, just identified*, etc.), because it is capable of perfectly fitting any data whatsoever.

Whereas Wright's method is partly graphical and partly algebraic, the theory of directed graphs permits us to analyze questions of testability and identifiability in purely graphical terms, prior to data collection, and it also enables us to extend these analyses from linear to nonlinear or nonparametric models. This section deals with issues of testability in linear and nonparametric models.

5.2.1 The Testable Implications of Structural Models

When we hypothesize a model of the data-generating process, that model often imposes restrictions on the statistics of the data collected. In observational studies, these restrictions provide the only view under which the hypothesized model can be tested or falsified. In many cases, such restrictions can be expressed in the form of zero partial correlations; more significantly, the restrictions are implied by the structure of the path diagram alone, independent of the numerical values of the parameters, as revealed by the d-separation criterion.

Preliminary Notation

Before addressing the testable implication of structural models, let us first review some definitions from Section 1.4 and relate them to the standard notation used in the SEM

literature.

The graphs we discuss in this chapter represent sets of structural equations of the form

$$x_i = f_i(pa_i, \epsilon_i,) \qquad i = 1, \ldots, n, \tag{5.1}$$

where pa_i (connoting *parents*) stands for the set of variables judged to be immediate causes of X_i and where the ϵ_i represent errors due to omitted factors. Equation (5.1) is a nonlinear, nonparametric generalization of the standard linear equations

$$x_i = \sum_{k \neq i} \alpha_{ik} x_k + \epsilon_i, \qquad i = 1, \ldots, n, \tag{5.2}$$

in which pa_i correspond to those variables on the r.h.s. of (5.2) that have nonzero coefficients. A set of equations in the form of (5.1) will be called a *causal model* if each equation represents the process by which the value (not merely the probability) of variable X_i is selected. The graph G obtained by drawing an arrow from every member of pa_i to X_i will be called a *causal diagram*. In addition to full arrows, a causal diagram should contain a bidirected (i.e. double-arrowed) arc between any pair of variables whose corresponding errors are dependent.

It is important to emphasize that causal diagrams (as well as traditional path diagrams) should be distinguished from the wide variety of graphical models in the statistical literature whose construction and interpretation rest solely on properties of the joint distribution (Kiiveri et al. 1984; Whittaker 1990; Cox and Wermuth 1996; Lauritzen 1996; Andersson et al. 1999). The missing links in those statistical models represent conditional independencies, whereas the missing links in causal diagrams represent absence of causal connections (see note 3 and Section 5.4), which may or may not imply conditional independencies in the distribution.

A causal model will be called *Markovian* if its graph contains no directed cycles and if its ϵ_i are mutually independent (i.e., if there are no bidirected arcs). A model is *semi-Markovian* if its graph is acyclic and if it contains dependent errors.

If the ϵ_i are multivariate normal (a common assumption in the SEM literature), then the X_i in (5.2) will also be multivariate normal and will be fully characterized by the correlation coefficients ρ_{ij}. A useful property of multivariate normal distributions is that the conditional variance $\sigma^2_{X|z}$, conditional covariance $\sigma_{XY|z}$, and conditional correlation coefficient $\rho_{XY|z}$, are all independent of the value z. These are known as *partial* variance, covariance, and correlation coefficient and are denoted by $\sigma_{X \cdot Z}$, $\sigma_{XY \cdot Z}$, and $\rho_{XY \cdot Z}$ (respectively), where X and Y are single variables and Z is a set of variables. Moreover, the partial correlation coefficient $\rho_{XY \cdot Z}$ is zero if and only if $(X \perp\!\!\!\perp Y | Z)$ holds in the distribution.

The *partial regression coefficient* is given by

$$r_{YX \cdot Z} = \rho_{YX \cdot Z} \frac{\sigma_{Y \cdot Z}}{\sigma_{X \cdot Z}};$$

it is equal to the coefficient of Y in the linear regression of Y on X and Z (the order of the subscripts is essential). In other words, the coefficient of x in the regression equation

$$y = ax + b_1 z_1 + \ldots + b_k z_k$$

is given by

$$a = r_{YX \cdot Z_1 Z_2 \ldots Z_k}.$$

These coefficients can therefore be estimated by the method of least squares (Crámer 1946).

d-Separation and Partial Correlations

Markovian models (the parallel term in the SEM literature is *recursive models*;[7] Bollen 1989) satisfy the Markov property of Theorem 1.2.7; as a result, the statistical parameters of Markovian models can be estimated by ordinary regression analysis. In particular, the d-separation criterion is valid in such models (here we restate Theorem 1.2.4).

Theorem 5.2.1 (Verma and Pearl 1988; Geiger et al. 1990)
If sets X and Y are d-separated by Z in a DAG G, then X is independent of Y conditional on Z in every Markovian model structured according to G. Conversely, if X and Y are not *d-separated by Z in a DAG G, then X and Y are dependent conditional on Z in* almost all *Markovian models structured according to G.*

Because conditional independence implies zero partial correlation, Theorem 5.2.1 translates into a graphical test for identifying those partial correlations that must vanish in the model.

Corollary 5.2.2 *In any Markovian model structured according to a DAG G, the partial correlation $\rho_{XY \cdot Z}$ vanishes whenever the nodes corresponding to the variables in Z d-separate node X from node Y in G, regardless of the model's parameters. Moreover, no other partial correlation would vanish for all the model's parameters.*

Unrestricted semi-Markovian models can always be emulated by Markovian models that include latent variables, with the latter accounting for all dependencies among error terms. Consequently, the d-separation criterion remains valid in such models if we interpret bidirected arcs as emanating from latent common parents. This may not be possible in some linear semi-Markovian models where each latent variable is restricted to influence at most two observed variables (Spirtes et al. 1996). However, it has been shown that the d-separation criterion remains valid in such restricted systems (Spirtes et al. 1996) and, moreover, that the validity is preserved when the network contains cycles (Spirtes et al. 1998; Koster 1999). These results are summarized in the next theorem.

Theorem 5.2.3 (d-Separation in General Linear Models)
For any linear model structured according to a diagram D, which may include cycles and bidirected arcs, the partial correlation $\rho_{XY \cdot Z}$ vanishes if the nodes corresponding to the set of variables Z d-separate node X from node Y in D. (Each bidirected arc $i \dashleftarrow\!\dashrightarrow j$ is interpreted as a latent common parent $i \leftarrow L \rightarrow j$.)

[7] The term *recursive* is ambiguous; some authors exclude correlated errors but others do not.

For linear structural equation models (see (5.2)), Theorem 5.2.3 implies that those (and only those) partial correlations identified by the d-separation test are guaranteed to vanish independent of the model parameters α_{ik} and independent of the error variances. This suggests a simple and direct method for testing models: rather than going through the standard exercise of finding a maximum likelihood estimate for the model's parameters and scoring those estimates for fit to the data, we can directly test for each zero partial correlation implied by the free model. The advantages of using such tests were noted by Shipley (1997), who also devised implementations of these tests.

However, the question arises of whether it is feasible to test for the vast number of zero partial correlations entailed by a given model. Fortunately, these partial correlations are not independent of each other; they can be derived from a relatively small number of partial correlations that constitutes a *basis* for the entire set (Pearl and Verma 1987).

Definition 5.2.4 (Basis)
Let S be a set of partial correlations. A basis *B for S is a set of zero partial correlations where* (i) *B implies (using the laws of probability) the zero of every element of S and* (ii) *no proper subset of B sustains such implication.*

An obvious choice of a basis for the zero partial correlations entailed by a DAG D is the set of equalities $B = \{\rho_{ij \cdot pa_i} = 0 | i > j\}$, where i ranges over all nodes in D and j ranges over all predecessors of i in any order that agrees with the arrows of D. In fact, this set of equalities reflects the "parent screening" property of Markovian models (Theorem 1.2.7), which is the source of all the probabilistic information encoded in a DAG . Testing for these equalities is therefore sufficient for testing all the statistical claims of a linear Markovian model. Moreover, when the parent sets PA_i are large, it may be possible to select a more economical basis, as shown in the next theorem.[8]

Theorem 5.2.5 (Graphical Basis)
Let (i, j) be a pair of nonadjacent nodes in a DAG D, and let Z_{ij} be any set of nodes that are closer to i than j is to i and such that Z_{ij} d-separates i from j. The set of zero partial correlations $B = \{\rho_{ij \cdot Z_{ij}} = 0 | i > j\}$, consisting of one element per nonadjacent pair, constitutes a basis for the set of all zero partial correlations entailed by D.

Theorem 5.2.5 states that the set of zero partial correlations corresponding to *any* separation between nonadjacent nodes in the diagram encapsulates all the statistical information conveyed by a linear Markovian model. A proof of Theorem 5.2.5 is given in Pearl and Meshkat (1999).

Examining Figure 5.1, we see that each of following two sets forms a basis for the model in the figure:

$$\begin{aligned} B_1 &= \{\rho_{32 \cdot 1} = 0,\ \rho_{41 \cdot 3} = 0,\ \rho_{42 \cdot 3} = 0,\ \rho_{51 \cdot 43} = 0,\ \rho_{52 \cdot 43} = 0\}, \\ B_2 &= \{\rho_{32 \cdot 1} = 0,\ \rho_{41 \cdot 3} = 0,\ \rho_{42 \cdot 1} = 0,\ \rho_{51 \cdot 3} = 0,\ \rho_{52 \cdot 1} = 0\}. \end{aligned} \tag{5.3}$$

The basis B_1 employs the parent set PA_i for separating i from j, $(i > j)$. Basis

[8] The possibility that linear models may possess more economical bases came to my awareness during a conversation with Rod McDonald.

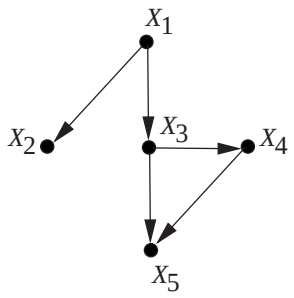

Figure 5.1: Model testable with two regressors for each missing link (equation (5.3)).

B_2, on the other hand, employs smaller separating sets and thus leads to tests involve fewer regressors. Note that each member of a basis corresponds to a missing arrow in the DAG; therefore, the number of tests required to validate a DAG is equal to the number of missing arrows it contains. The sparser the graph, the more it constrains the covariance matrix and more tests are required to verify those constraints.

5.2.2 Testing the Testable

In linear structural equation models, the hypothesized causal relationships between variables can be expressed in the form of a directed graph annotated with coefficients, some fixed a priori (usually to zero) and some free to vary. The conventional method for testing such a model against the data involves two stages. First, the free parameters are estimated by iteratively maximizing a fitness measure such as the likelihood function. Second, the covariance matrix implied by the estimated parameters is compared to the sample covariances and a statistical test is applied to decide whether the latter could originate from the former (Bollen 1989; Chou and Bentler 1995).

There are two major weaknesses to this approach:

1. if some parameters are not identifiable, then the first phase may fail to reach stable estimates for the parameters and the investigator must simply abandon the test;

2. if the model fails to pass the data fitness test, the investigator receives very little guidance about which modeling assumptions are wrong.

For example, Figure 5.2 shows a path model in which the parameter α is not identifiable if $\mathrm{cov}(\epsilon_1, \epsilon_2)$ is assumed to be unknown, which means that the maximum likelihood method may fail to find a suitable estimate for α, thus precluding the second phase of the test. Still, this model is no less testable than the one in which $\mathrm{cov}(\epsilon_1, \epsilon_2) = 0$, α is identifiable, and the test can proceed. These models impose the same restrictions on the covariance matrix—namely, that the partial correlation $\rho_{XZ.Y}$ should vanish (i.e., $\rho_{XZ} = \rho_{XY}\rho_{YZ}$)—yet the model with free $\mathrm{cov}(\epsilon_1, \epsilon_2)$, by virtue of α being nonidentifiable, cannot be tested for this restriction.

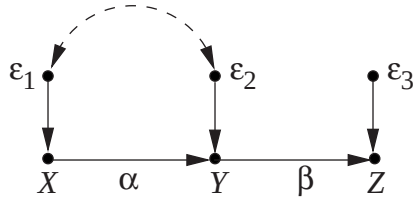

Figure 5.2: A testable model containing unidentified parameter (α).

Figure 5.3 illustrates the weakness associated with model diagnosis. Suppose the true data-generating model has a direct causal connection between X and W, as shown in Figure 5.3(a), while the hypothesized model (Figure 5.3(b)) has no such connection. Statistically, the two models differ in the term $\rho_{XW \cdot Z}$, which should vanish according to Figure 5.3(b) and is left free according to Figure 5.3(a). Once the nature of the

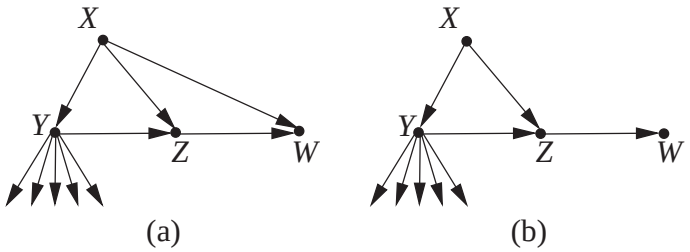

Figure 5.3: Models differing in one local test, $\rho_{XW \cdot Z} = 0$.

discrepancy is clear, the investigator must decide whether substantive knowledge justifies alteration of the model by adding either a link or a curved arc between X and W. However, because the effect of the discrepancy will be spread over several covariance terms, global fitness tests will not be able to isolate the discrepancy easily. Even multiple fitness tests on various local modifications of the model (such tests are provided by LISREL) may not help much, because the results may be skewed by other discrepancies in different parts of the model, such as the subgraph rooted at Y. Thus, testing for global fitness is often of only minor use in model debugging.

An attractive alternative to global fitness testing is local fitness testing, which involves listing the restrictions implied by the model and testing them one by one. A restriction such as $\rho_{XW \cdot Z} = 0$, for example, can be tested locally without measuring Y or any of its descendants, thus keeping errors associated with those measurements from interfering with the test for $\rho_{XW \cdot Z} = 0$, which is the real source of the lack of fit. More generally, typical SEM models are often close to being "saturated," claiming but a few restrictions in the form of a few edges missing from large, otherwise unrestrictive diagrams. Local and direct tests for those restrictions are more reliable than global tests, since they involve fewer degrees of freedom and are not contaminated with irrelevant measurement errors. The missing edges approach described in Section

5.2.1 provides a systematic way of detecting and enumerating the local tests needed for testing a given model.

5.2.3 Model Equivalence

In Section 2.3 (Definition 2.3.3) we defined two structural equation models to be observationally equivalent if every probability distribution that is generated by one of the models can also be generated by the other. In standard SEM, models are assumed to be linear and data are characterized by covariance matrices. Thus, two such models are observationally indistinguishable if they are *covariance equivalent*, that is, if every covariance matrix generated by one model (through some choice of parameters) can also be generated by the other. It can be easily verified that the equivalence criterion of Theorem 1.2.8 extends to covariance equivalence.

Theorem 5.2.6 *Two Markovian linear-normal models are covariance equivalent if and only if they entail the same sets of zero partial correlations. Moreover, two such models are covariance equivalent if and only if their corresponding graphs have the same sets of edges and the same sets of v-structures.*

The first part of Theorem 5.2.6 defines the testable implications of Markovian models. It states that, in nonmanipulative studies, Markovian structural equation models cannot be tested for any feature other than those zero partial correlations that the d-separation test reveals. It also provides a simple test for equivalence that requires, instead of checking all the d-separation conditions, merely a comparison of corresponding edges and their directionalities.

In semi-Markovian models (DAGs with correlated errors), the d-separation criterion is still valid for testing independencies (see Theorem 5.2.3), but independence equivalence no longer implies observational equivalence.[9] Two models that entail the same set of zero partial correlations among the observed variables may yet impose different inequality constraints on the covariance matrix. Nevertheless, Theorems 5.2.3 and 5.2.6 still provide necessary conditions for testing equivalence.

Generating Equivalent Models

By permitting arrows to be reversed as long as no v-structures are destroyed or created, we can use Theorem 5.2.6 to generate equivalent alternatives to any Markovian model. Meek (1995) and Chickering (1995) showed that $X \rightarrow Y$ can be replaced by $X \leftarrow Y$ if and only if all parents of X are also parents of Y. They also showed that, for any two equivalent models, there is always some sequence of such edge reversals that takes one model into the other. This simple rule for edge reversal coincides with those proposed by Stelzl (1986) and Lee and Hershberger (1990).

In semi-Markovian models, the rules for generating equivalent models are more complicated. Nevertheless, Theorem 5.2.6 yields convenient graphical principles for testing the correctness of edge-replacement rules. The basic principle is that if

[9] Verma and Pearl (1990) presented an example using a nonparametric model, and Richardson devised an example using linear models with correlated errors (Spirtes and Richardson 1996).

we regard each bidirected arc $X \dashleftarrow\!\!\dashrightarrow Y$ as representing a latent common cause $X \leftarrow L \rightarrow Y$, then the "if" part of Theorem 5.2.6 remains valid; that is, any edge-replacement transformation that does not destroy or create a v-structure is allowed. Thus, for example, an edge $X \rightarrow Y$ can be replaced by a bidirected arc $X \dashleftarrow\!\!\dashrightarrow Y$ whenever X and Y have no other parents, latent or observed. Likewise, an edge $X \rightarrow Y$ can be replaced by a bidirected arc $X \dashleftarrow\!\!\dashrightarrow Y$ whenever (1) X and Y have no latent parents and (2) every parent of X or Y is a parent of both. Such replacements do not introduce new v-structures. However, since v-structures may now involve latent variables, we can tolerate the creation or destruction of some v-structures as long as this does not affect partial correlations among the observed variables. Figure 5.4(a) demonstrates that the creation of certain v-structures can be tolerated. By reversing the arrow $X \rightarrow Y$ we create two converging arrows $Z \rightarrow X \leftarrow Y$ whose tails are connected, not directly, but through a latent common cause. This is tolerated because, although the new convergence at X blocks the path (Z, X, Y), the connection between Z and Y (through the arc $Z \dashleftarrow\!\!\dashrightarrow Y$) remains unblocked and, in fact, cannot be blocked by any set of observed variables.

We can carry this principle further by generalizing the concept of v-structure. Whereas in Markovian models a v-structure is defined as two converging arrows whose tails are not connected by a link, we now define v-structure as any two converging arrowheads whose tails are "separable." By *separable* we mean that there exists a conditioning set S capable of d-separating the two tails. Clearly, the two tails will not be separable if they are connected by an arrow or by a bidirected arc. But a pair of nodes in a semi-Markovian model can be inseparable even when not connected by an edge (Verma and Pearl 1990). With this generalization in mind, we can state necessary conditions for edge replacement as follows.

Rule 1: An arrow $X \rightarrow Y$ is interchangeable with $X \dashleftarrow\!\!\dashrightarrow Y$ only if every neighbor or parent of X is inseparable from Y. (By *neighbor* we mean a node connected (to X) through a bidirected arc.)

Rule 2: An arrow $X \rightarrow Y$ can be reversed into $X \leftarrow Y$ only if, before reversal, (i) every neighbor or parent of Y (excluding X) is inseparable from X and (ii) every neighbor or parent of X is inseparable from Y.

For example, consider the model $Z \dashleftarrow\!\!\dashrightarrow X \rightarrow Y$. The arrow $X \rightarrow Y$ cannot be replaced with a bidirected arc $X \dashleftarrow\!\!\dashrightarrow Y$ because Z (a neighbor of X) is separable from Y by the set $S = \{X\}$. Indeed, the new v-structure created at X would render X and Y marginally independent, contrary to the original model.

As another example, consider the graph in Figure 5.4(a). Here, it is legitimate to replace $X \rightarrow Y$ with $X \dashleftarrow\!\!\dashrightarrow Y$ or with a reversed arrow $X \leftarrow Y$ because X has no neighbors and Z, the only parent of X, is inseparable from Y. The same considerations apply to Figure 5.4(b); variables Z and Y, though nonadjacent, are inseparable, because the paths going from Z to Y through W cannot be blocked.

A more complicated example, one that demonstrates that rules 1 and 2 are not sufficient to ensure the legitimacy of a transformation, is shown in Figure 5.4(c). Here, it appears that replacing $X \rightarrow Y$ with $X \dashleftarrow\!\!\dashrightarrow Y$ would be legitimate because the (latent) v-structure at X is shunted by the arrow $Z \rightarrow Y$. However, the original model

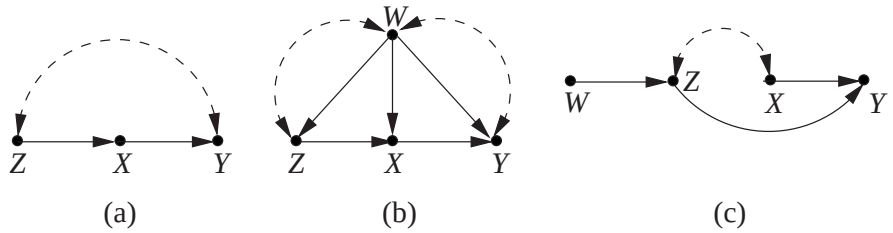

Figure 5.4: Models permitting ((a) *and* (b)) and forbidding (c) the reversal of XY.

shows the path from W to Y to be d-connected given Z, whereas the postreplacement model shows the same path d-separated given Z. Consequently, the partial correlation $\rho_{WY \cdot Z}$ vanishes in the postreplacement model but not in the prereplacement model. A similar disparity also occurs relative to the partial correlation $\rho_{WY \cdot ZX}$. The original model shows that the path from W to Y is blocked, given $\{Z, X\}$, but the postreplacement model shows that path to be d-connected, given $\{Z, X\}$. Consequently, the partial correlation $\rho_{WY \cdot ZX}$ vanishes in the prereplacement model but is unconstrained in the postreplacement model.[10] Evidently, it is not enough to impose rules on the parents and neighbors of X; remote ancestors (e.g. W) should be considered, too.

These rules are just a few of the implications of the d-separation criterion when applied to semi-Markovian models. A necessary and sufficient criterion for testing the d-separation equivalence of two semi-Markovian models was devised by Spirtes and Verma (1992). Spirtes and Richardson (1996) extended that criterion to include models with feedback cycles. However, we should keep in mind that, because two semi-Markovian models can be zero-partial-correlation equivalent and yet not covariance equivalent, criteria based on d-separation can provide merely the necessary conditions for model equivalence.

The Significance of Equivalent Models

Theorem 5.2.6 is methodologically significant because it clarifies what it means to claim that structural models are "testable" (Bollen 1989, p. 78).[11] It asserts that we never test *a* model but rather a whole *class* of observationally equivalent models from which the hypothesized model cannot be distinguished by any statistical means. It asserts as well that this equivalence class can be constructed (by inspection) from the graph, which thus provides the investigator with a vivid representation of competing alternatives for consideration. Graphs representing all models in a given equivalence class have been devised by Verma and Pearl (1990) (see Section 2.6), Spirtes et al. (1993), and Andersson et al. (1999). Richardson (1996) discusses the representation of equivalence classes of models with cycles.

[10] This example was brought to my attention by Jin Tian, and a similar one by two anonymous reviewers.

[11] In response to an allegation that "path analysis does not derive the causal theory from the data, or test any major part of it against the data" (Freedman 1987, p. 112), Bollen (1989, p. 78) stated, "we can test and reject structural models.... Thus the assertion that these models cannot be falsified has little basis."

Although it is true that (overidentified) structural equation models have testable implications, those implications are but a small part of what the model represents: a set of claims, assumptions, and implications. Failure to distinguish among causal assumptions, statistical implications, and policy claims has been one of the main reasons for the suspicion and confusion surrounding quantitative methods in the social sciences (Freedman 1987, p. 112; Goldberger 1992; Wermuth 1992). However, because they make the distinctions among these components vivid and crisp, graphical methods promise to make SEM more acceptable to researchers from a wide variety of disciplines.

By and large, the SEM literature has ignored the explicit analysis of equivalent models. Breckler (1990), for example, found that only one of 72 articles in the areas of social and personality psychology even acknowledged the existence of an equivalent model. The general attitude has been that the combination of data fitness and model over-identification is sufficient to confirm the hypothesized model. Recently, however, the existence of multiple equivalent models seems to have jangled the nerves of some SEM researchers. MacCallum et al. (1993, p. 198) concluded that "the phenomenon of equivalent models represents a serious problem for empirical researchers using CSM" and "a threat to the validity of interpretation of CSM results" (CSM denotes "cxovariance structure modeling"; this does not differ from SEM, but the term is used by some social scientists to disquise euphemistically the causal content of their models). Breckler (1990, p. 262) reckoned that "if one model is supported, so too are all of its equivalent models" and hence ventured that "the term *causal modeling* is a misnomer."

Such extremes are not justifiable. The existence of equivalent models is logically inevitable if we accept the fact that causal relations cannot be inferred from statistical data alone; as Wright (1921) stated, "prior knowledge of the causal relations is assumed as prerequisite" in SEM. But this does not make SEM useless as a tool for causal modeling. The move from the qualitative causal premises represented by the structure of a path diagram (see note 3) to the quantitative causal conclusions advertised by the coefficients in the diagram is neither useless nor trivial. Consider, for example, the model depicted in Figure 5.5, which Bagozzi and Burnkrant (1979) used to illustrate problems associated with equivalent models. Although this model is saturated (i.e., just identified) and although it has (at least) 27 semi-Markovian equivalent models, finding that the influence of AFFECT on BEHAVIOR is almost three times stronger (on a standardized scale) than the influence of COGNITION on BEHAVIOR is still very illuminating—it tells us about the relative effectiveness of different behavior modification policies if some are known to influence AFFECT and others COGNITION. The significance of this quantitative analysis on policy analysis may be more dramatic when a path coefficient turns negative while the corresponding correlation coefficient measures positive. Such quantitative results may have profound impact on policy decisions, and learning that these results are logically implied by the data and the qualitative premises embedded in the diagram should make the basis for policy decisions more transparent to defend or to criticize.

In summary, social scientists need not abandon SEM altogether; they need only abandon the notion that SEM is a method of *testing* causal models. Structural equation modeling is a method of testing a tiny fraction of the premises that make up a causal model and, in cases where that fraction is found to be compatible with the data, the

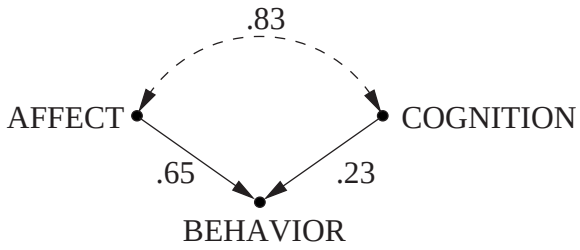

Figure 5.5: Untestable model displaying quantitative causal information derived.

method elucidates the necessary quantitative consequences of both the premises and the data. It follows, then, that users of SEM should concentrate on examining the implicit theoretical premises that enter into a model. As we will see in Section 5.4, graphical methods make these premises vivid and precise.