GRAPHICAL TOOLS FOR LINEAR PATH MODELS

Bryant Chen

IBM RESEARCH

Judea Pearl

UNIVERSITY OF CALIFORNIA, LOS ANGELES

Rex Kline

CONCORDIA UNIVERSITY, MONTREAL, CANADA

April 5, 2018

Correspondence should be sent to
E-Mail: bryant.chen@ibm.com, judea@cs.ucla.edu, rex.kline@concordia.edu
Phone: (310) 825-3243

Abstract

Graphical tools now play a pivotal role in the analysis of causal relationships in many data-intensive sciences, especially biostatistics, epidemiology, social science, and machine learning. While originally developed primarily for non-parametric models, many of these tools can also be applied to answer key research questions in linear path models, particularly questions of identification and model fit. In this paper, we survey graphical tools that enable psychology researchers to 1) identify path coefficients using partial regression, (2) identify path coefficients using instrumental variables, (3) apply local fit tests using vanishing partial correlation, and (4) apply local fit tests using overidentified model parameters. We show how these tasks can be accomplished in a qualitative manner, prior to taking any data, and with minimal arithmetic operations.

Key words: structural equation models; path models; graphical models; d-separation; local fit; goodness of fit; identification; regression; instrumental variables, equivalent models

1. Introduction

Recent advances in graphical models have had a transformative impact on causal analysis in almost every data-intensive discipline, especially biostatistics (Pearl, 1995), epidemiology (VanderWeele, 2015), machine learning (Koller and Friedman, 2009), and social science (Morgan and Winship, 2007; Elwert, 2013). Among the tasks facilitated by graphical models are: assessing model compatibility with data, identification, policy evaluation, bias control, mediation, external validity, and the analysis of counterfactuals and missing data (Pearl, 2014). Only a meager portion of these developments have found their way to mainstream behavioral science which, by and large, prefers algebraic over graphical representations (Bollen, 1989; Mulaik, 2009; Hoyle, 2012). One of the reasons for this disparity rests on the fact that modern graphical techniques were developed for non-parametric analysis, while structural equation models (SEM) are typically analyzed within the confines of linear-normal models, to which matrix algebra and powerful statistical tests are applicable.

The purpose of this survey is to introduce modern tools of graphical models to researchers in multivariate experimental psychology, and to describe the benefits and insights that these models can provide in the context of linear path models[1]. We will begin by introducing the basic notation and definitions used in path diagrams, including Wright's path tracing rules, and then introduce the more advanced notion of graph separation. This notion, which was originally developed for non-parametric analysis, has simple and meaningful interpretation in linear models in terms of vanishing partial correlations. Graph separation provides the basis for parameter identification, determining model equivalence, the detection of instrumental variables and the assessment of model testability.

These tasks, which will be introduced and discussed separately in subsequent sections of the paper, will be shown to be executable directly at the level of the qualitative graph structures, prior to obtaining any data. For example, a simple inspection of the graph structure tells researchers which path coefficients can be estimated using OLS, and which ones require 2SLS, and how to select appropriate instrumental variables from the model. These graphical criteria, therefore, provide valuable tools during the model building portion of the analysis, and facilitate meaningful communication among researchers from diverse disciplines.

Lastly, we demonstrate, using a real data set and freely available computer software, how these tools can be incorporated into and enrich research in mainstream structural equation modeling.

2. Path Diagrams and Graphs

In this survey, we consider path diagrams that are linear in their parameters and in which each node represents a substantive random variable present in the data. The value taken by each

---

[1]Contrasting other didactic papers on graphical methods (Pearl, 1995, 1998, 2012) this paper is written directly in a language familiar to applied researchers in psychology, with an emphasis on how graphical techniques can enrich standard modeling practices. Philosophical questions concerning the meaning of causation, the explicit analysis of interventions and counterfactuals, and generalizations to non-parametric structural models are deferred to the cited references.
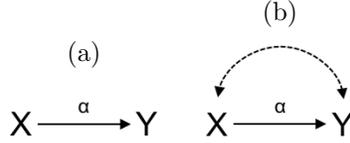
Figure 1: (a) α is identified (b) α is not identified

variable is determined by some of its adjacent variables plus latent random errors, which may be correlated. The target of analysis is the estimation of the path coefficients and the degree to which the model fits the available data. Latent variable models, such as confirmatory factor analysis (CFA) models, are not the focus of this paper, though many of the graphical tools described can also be applied to such models (Chen and Pearl, 2014; Thoemmes et al., 2018).

Graphical methods typically focus on the implications that the model's structure has on the probability distibution over the modeled variables. Whenever two variables in a path diagram are *not* connected by a direct path, the model restricts the corresponding path coefficients to zero. Similarly, whenever two variables in a path diagram are *not* connected by a bidirected path, the model restricts the covariance of the corresponding error terms to zero[2]. These restrictions have implications on the probability distribution over the modeled variables. For example, in the simplest case, consider a two variable model, where the two variables are not connected by any paths whatsoever. This model implies that the two variables are uncorrelated and independent, and this implication can be used to test the model and assess goodness of fit. Namely, we can test the validity of this two-variable model by checking whether the variables are actually uncorrelated in the data. If they are, in fact, correlated, then the model does not fit the data and must be misspecified.

Another possible implication of the model's restrictions is that a path coefficient is always equal to some function of the covariance matrix. In this case, we say the coefficient is *identified*. If the model does not sufficiently restrict the value of a coefficient and multiple values are consistent with the implied probability distribution, then the coefficient is *not identified*, and it cannot be estimated from data. Identification will be discussed in more detail in subsequent sections. For example, the simple path diagram shown in Figure 1a represents a single structural equation, $Y = \alpha X + U_Y$, with U_Y uncorrelated with X. Assuming without loss of generality that the variables are standardized to mean 0 and variance 1, the covariance of X and Y, $\sigma(X, Y) = \sigma(X, \alpha X + U_Y) = \alpha \sigma(X, X) + \sigma(X, U_Y) = \alpha$. Thus, we see that the lack of a bidirected path between X and Y implies that α is equal to the covariance of X and Y. In contrast, if there were a bidirected path between X and Y, then $\sigma(X, U_Y)$ would be an unknown quantity, and all possible values of α are compatible with $\sigma(X, Y)$. Thus, α cannot be estimated from the sample covariance matrix. In this paper, we survey basic graphical methods for deriving identifiability of path coefficients and testable implications.

---

[2]Note that we do not use term "bidirected" to denote reciprocal relations between variables. In this paper, bidirected paths signify correlated error terms.
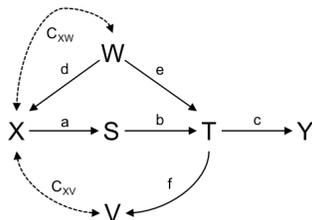
Figure 2: Model illustrating Wright's path tracing rules and d-separation

2.1. Nomenclature

In this subsection, we "translate" from terms and symbols in graph representations to counterparts more recognizable to psychology researchers. Both graphical models and the path diagrams introduced by Wright (1921) are meant to convey the causal relations among the variables in the data set[3]. In both, directed arrows signify direct causal relations between two variables. However, the usage of bidirected arrows differs slightly. Wright used bidirected arrows to denote non-zero correlation between exogenous variables. In graphical models, a bidirected arrow signifies a latent common cause between the two variables or, equivalently, correlated error terms in the corresponding structural equations[4]. To simplify the usage of d-separation, which will be introduced in a subsequent section, we will use the graphical modeling convention, where bidirected arrows signify correlated error terms.

In the graphical modeling literature, variables in the graph are often called nodes or vertices, and paths are called edges or links. A "path" in the graphical modeling literature refers instead to a sequence of contiguous edges and nodes that does not involve the same node twice. (A single edge and the two nodes connecting it also qualifies as a path.) To avoid confusion, we will adopt the nomenclature more familiar to psychology researchers, where a path refers to a single arrow or bidirected arrow. A sequence of one or more contiguous paths and their corresponding variables (that does not involve the same variable twice) will instead be called a *route*. Together, the path coefficients associated with directed paths and the error covariances associated with bidirected paths are the model *parameters*.

A route may go either along or against the direction of the arrows. A *directed route* from variable X to variable Y is one that consists of only directed arrows from A to F. For example, the route $X \to S \to T \to Y$ in Figure 2 is a directed route. A *collider route* between X and Y is one that contains "colliding" arrowheads that both point to the same variable. For example, $X \leftarrow\!\dashrightarrow V \leftarrow T \to Y$ is a collider route between X and Y due to the collision at V. In this context, we call V a *collider*, noting that V may or may not be a collider depending on the route in

---

[3]Some users of path diagrams prefer not to endow the model with causal meaning. We will follow Sewall Wright's causal intrepretation and causal terminology for ease of communication and explanation. Most of our identification and goodness of fit techniques are equally valid for alternative interpretations of path coefficients.

[4]Since we do not consider path coefficients between latent variables, the presence of latent variables are taken into account through the correlations they induce on the error terms (Pearl, 2009).

question. For example, if there was another route $X \to V \to Y$, then V would not be a collider in the context of this route. When we introduce Wright's rules and d-separation, we will see that collider routes, unlike other routes, do not transmit correlation between variables. Lastly, a *back-door route* between X and Y is one that begins with an arrow pointed at X and ends with an arrow pointed at Y but does not contain a collider. For example, both $X \gets W \to T \to Y$ and $X \gets\!\!\dashrightarrow W \to T \to Y$ are back-door routes.

We will call a diagram or model *recursive* if it does not contain any *cycles*, that is a directed route that begins and ends with the same variable, and it does not contain any bidirected paths[5]. Models that are not recursive are called non-recursive, while non-recursive models without cycles are additionally called *partially recursive*. For example, Figure 2 is both non-recursive and partially recursive because it contains a bidirected path, but no cycles. Lastly, unless otherwise stated, we will assume without loss of generality that all variables have been standardized to mean 0 and variance 1.

2.2. Wright's Path Tracing Rules

Wright introduced the path tracing rules in his paper, "Correlation and Causation" (Wright, 1921). These simple rules allow researchers to express the model-implied covariance between any two variables in terms of the model parameters. As a result, they are a valuable tool when analyzing the identifiability of parameters.

Wright's rules state that the covariance, $\sigma(Y, X)$, between any pair of variables, X and Y, is equal to the sum of products of path coefficients and error covariances along non-collider routes between X and Y. Formally, let $\Pi = \{\pi_1, \pi_2, ..., \pi_k\}$ denote the non-collider routes between X and Y, and let p_i be the product of path coefficients along route π_i. Then the model-implied covariance between variables X and Y is $\sum_i p_i$. For example, the covariance between A and F implied by Figure 2 is equal to $C_{XW} \cdot e \cdot c + d \cdot e \cdot c + a \cdot b \cdot c$, where C_{XW} is the covariance between the error terms of X and W. The first term corresponds to the route $X \gets\!\!\dashrightarrow W \to T \to Y$, the second term corresponds to $X \gets W \to T \to Y$, and the last to $X \to S \to T \to Y$. The route, $X \gets\!\!\dashrightarrow V \gets T \to Y$ is not included because it contains a collider.

To express the partial covariance, $\sigma(Y, X|Z)$, partial correlation, $\rho(Y, X|Z)$ or regression coefficient, $\beta(Y, X|Z)$, of Y on X given Z in terms of structural coefficients we can first apply the following reductions given by Crámer (1946), before utilizing Wright's rules. When Z is a single variable, as opposed to a set, these reductions are:

---

[5]In the graphical modeling literature, the term "recursive" is often used to describe models that do not contain cycles but may or may not contain bidirected paths. To avoid confusion, we adopt the terminology more familiar to psychology researchers, where a recursive models is one in which there are no cycles or bidirected paths.
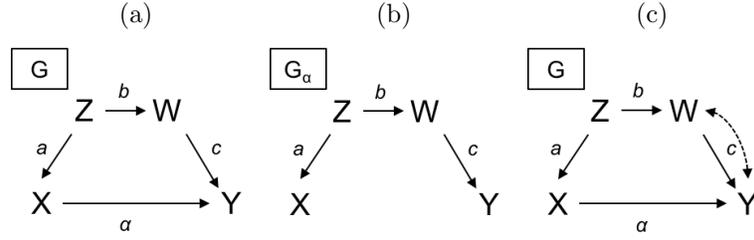
Figure 3: (a) α is identified by adjusting for Z or W (b) The graph G_α used in the identification of α (c) α is identified by adjusting for Z (or Z and W) but not W alone

$$\rho(Y,X|Z) = \frac{\rho(Y,X) - \rho(Y,Z)\rho(X,Z)}{[(1 - \rho^2(Y,Z))(1 - \rho^2(X,Z))]^{\frac{1}{2}}} \tag{1}$$

$$\sigma(Y,X|Z) = \sigma(Y,X) - \frac{\sigma(Y,Z)\sigma(Z,X)}{\sigma(Z,Z)} \tag{2}$$

$$\beta(Y,X|Z) = \frac{\sigma(Y)}{\sigma(X)} \frac{\rho(Y,X) - \rho(Y,Z)\rho(Z,X)}{1 - \rho(X,Z)^2} \tag{3}$$

When Z is a single variable and S a set, we can reduce $\rho(Y,X|Z,S)$, $\sigma(Y,X|Z,S)$, or $\beta(Y,X|Z,S)$ as follows:

$$\rho(Y,X|Z,S) = \frac{\rho(Y,X|S) - \rho(Y,Z|S)\rho(X,Z|S)}{[(1 - \rho(Y,Z|S)^2)(1 - \rho(X,Z|S)^2)]^{\frac{1}{2}}} \tag{4}$$

$$\sigma(Y,X|Z,S) = \sigma_(Y,X|S) - \frac{\sigma(Y,Z|S)\sigma(Z,X|S)}{\sigma(Z|S)^2} \tag{5}$$

$$\beta(Y,X|Z,S) = \frac{\sigma(Y|S)}{\sigma(X|S)} \frac{\rho(Y,X|S) - \rho(Y,Z|S)\rho(Z,X|S)}{1 - \rho(X,Z|S)^2} \tag{6}$$

We see that $\rho(Y,X|Z,S)$, $\sigma(Y,X|Z,S)$, or $\beta(Y,X|Z,S)$ can be expressed in terms of pair-wise coefficients by recursively applying the above formulas for each element of S. Then, using Equations 1-6, we can express the reduced pairwise covariances / correlations in terms of
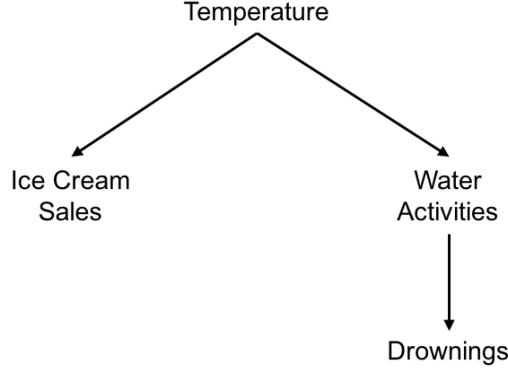
Figure 4: Diagram illustrating why Ice Cream Sales and Drowning are uncorrelated given Temperature and/or Water Activities

the structural coefficients. For example, reducing $\beta(Y, X|W)$ for Figure 3a can be done as follows:

$$\beta(Y, X|W) = \frac{\sigma(Y)}{\sigma(X)} \frac{\rho(Y,X) - \rho(Y,W)\rho(W,X)}{1 - \rho(X,W)^2} \tag{7}$$

$$= \frac{1}{1} \frac{(\alpha + abc) - (c + ba\alpha)(ab)}{1 - a^2 b^2} \tag{8}$$

$$= \frac{\alpha + abc - abc - a^2 b^2 \alpha}{1 - a^2 b^2} \tag{9}$$

$$= \frac{\alpha(1 - a^2 b^2)}{1 - a^2 b^2} \tag{10}$$

$$= a \tag{11}$$

Eq. 11 implies that the path coefficient a is identifiable and can be estimated by OLS using the partial regression of Y on X and W. In theory, the identifiability of any parameter in a linear-normal model can be determined by analyzing the equations derived using Wright's rules. However, the computation required is often tedious, and sociologists in the 1960's struggled to avoid it (Alwin and Hauser, 1975; Wolfle, 1980). In section Identification, we will show that Equation 11 can be written by inspection, given the structure of the model in Figure 3a.

2.3. D-Separation

The importance of deriving model-implied zero partial correlations in assessing goodness of fit was recognized by social scientists as early as the 1960s. Blalock (1962) even provided an exhaustive enumeration of the implied vanishing partial correlations for all possible recursive, four-variable models. In this section, we introduce *d-separation*, a graphical technique for deriving model-implied conditional independences and zero partial correlations by analyzing the path
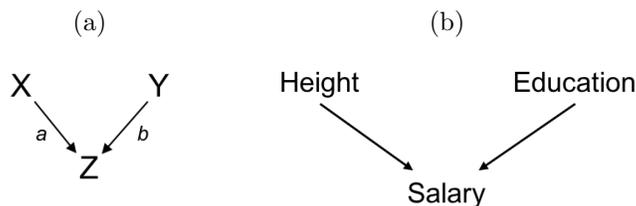
Figure 5: Examples illustrating conditioning on a collider

diagram[6]. Not only will this technique allow us to derive testable implications of the model for assessing goodness of fit, but it will also be utilized extensively in the analysis of identification that follows.

Wright's rules show that certain routes induce association between variables, while others do not. (We will call routes that induce association in this manner *active*.) Conditioning on other variables can change the nature of these routes. Active routes may become inactive, while inactive routes may become active. In this section, we describe three rules that govern whether a given route is active or inactive, given a set of conditioning variables. These three rules will be used to determine whether the model implies that variables are conditionally independent given a conditioning set or *d-separated*. The first rule follows from Wright's rules.

Rule 1: When no variables are conditioned upon, a route between two variables is active if and only if it does not contain a collider.

The dependency and information transmitted along an active, non-collider route can be blocked and the route rendered inactive by conditioning on variables along the route. For example, consider the path diagram shown in Figure 4, which models the correlation between ice cream sales and drownings. When the weather gets warm, people tend to both buy ice cream and play in the water, resulting in both increased ice cream sales and drowning deaths. In the language of graphical models, ice cream sales and drowning are correlated due to the active route, Ice Cream Sales ← Temperature → Water Activities → Drownings.

However, if we only consider days with the same temperature and/or the same number of people engaging in water activities, then the correlation between Ice Cream Sales and Drownings will vanish. Thus, conditioning on either Temperature or Ice Cream Sales *blocks* the flow of correlation transmitted through the route. In general, conditioning on non-colliders along a route blocks and renders it inactive.

Rule 2: An active route between two variables, X and Y, is rendered inactive by conditioning on a non-collider along that route.

---

[6]See also Hayduk et al. (2003), (Mulaik, 2009, ch. 4), and (Kline, 2016, ch. 8) for an introduction to d-separation tailored to SEM practitioners.

From Wright's rules, we know that a collider also blocks the flow of information along a route and renders it inactive. However, when a collider or one of its descendants is conditioned upon and taken as given, that variable no longer blocks any route for which it is a collider. This is due to the fact that conditioning on a collider or its descendant opens the flow of information between the parents of the collider. For example, X and Y are uncorrelated in Figure 5a. However, conditioning on the collider, Z, correlates X and Y giving $\rho(X, Y|Z) \neq 0$. This phenomenon is known Berkson's paradox or "explaining away".

To illustrate, consider the example depicted in Figure 5b. It is common knowledge that higher education often affords one a greater salary. Additionally, studies have shown that height also has a positive impact on one's salary. Let us assume that there are no other determinants of salary and that Height and Education are uncorrelated. If we observe an individual with a high salary that is also short, the probability that the individual is highly educated increases. As a result, we see that observing Salary correlates Education and Height. Similarly, observing an effect or indicator of salary, say the individual's Ferrari, also correlates Education and Height.

The fact that $\sigma(Y, X|Z) \neq 0$ when $\sigma(Y, X) = 0$ and Z a common child of X and Y can also be illustrated using Wright's path tracing rules. Consider Figure 5a where Z is a common effect of X and Y. We have $\sigma(Y, X) = 0$ and, using Equation 2,

$$\sigma(Y, X|Z) = \sigma(Y, X) - \frac{\sigma(Y, Z)\sigma(Z, X)}{\sigma(Z, Z)}$$
$$= 0 - \frac{ab}{1}$$
$$= -ab.$$

When a and b are non-zero we have an algebraic confirmation of our intuition from the salary example that X and Y are uncorrelated marginally, but becoming correlated when we condition on Z.

Rule 3: If a collider is a member of the conditioning set Z, or has a descendant in Z, then the collider no longer blocks any route that traces it.

Rule 3 implies that routes containing colliders can become active by conditioning on colliders or their descendants. However, if a non-collider is also a member of the conditioning set, then Rule 2 implies that the route is still blocked by the non-collider and it is inactive.

Definition 1. We say that variables X and Y are *d-separated* given a conditioning set Z if there are no active routes between X and Y given Z.

Theorem 1. If X and Y are d-separated given a set Z in the path diagram, G, then G implies that X and Y are independent given Z and $\sigma(X, Y|Z) = \rho(X, Y|Z) = \beta(X, Y|Z) = \beta(YX|Z) = 0$.

If X and Y are d-connected given Z then $\sigma(X, Y|Z)$ is generally not equal to zero but may

equal zero for particular parameter values. For example, it is possible that the values of the coefficients are such that the active routes between X and Y perfectly cancel one another.

We use the diagram depicted in Figure 2 as an example to illustrate the rules of d-separation. In this example, X is d-separated from Y by $\{W, S\}$. Similarly X is d-separated from Y by $\{T\}$. However, X is not d-separated from Y by $\{W, S, V\}$ since conditioning on V opens the collider route $X \leftarrow\!\!\rightarrow V \leftarrow T \rightarrow Y$. Finally, V is not d-separated from W by conditioning on T since T is a descendant of X, opening the collider routes $W \leftarrow\!\!\rightarrow X \leftarrow\!\!\rightarrow V$ and $W \rightarrow X \leftarrow\!\!\rightarrow V$.

The concept of d-separation formalizes the intuition that routes carry associational information between variables and that this flow of information can be blocked by conditioning. This intuition drives many of the results in identification and assessing goodness of fit that will be discussed in subsequent sections, making it an essential component of graphical modeling.

We conclude this section by noting that d-separation can be used to derive vanishing partial correlations in both recursive and non-recursive models (Spirtes, 1995). Further, all vanishing partial correlations implied by a path model can be obtained using d-separation (Pearl, 2009, ch. 1.2.3). Finally, in fully recursive models where the variables are Gaussian, these vanishing partial correlations represent *all* of the model's testable implications, a point elaborated later.

3. Identification

Earlier, we discussed how a parameter is identified if the model implies that it is equal to some function of the implied probability distribution (e.g. covariance matrix). When identified, estimating the necessary aspects of the distribution (typically the sample covariance matrix in linear models) yields an estimate of the identified parameter. In contrast, if the model does not restrict the parameter to a single value in terms of the distribution, then it is not identified. In other words, the parameter is not identified if the model implies that it can be equal to multiple (often infinite) values in terms of the distribution. Clearly, any non-identified parameter cannot be estimated from data. When all of the model parameters are identified, then we say the model is identified. If a single parameter is not identified, we say the model is not identified.

In practice, a model's parameters are often estimated by submitting the data and model specification to a computer program, which attempts to find the parameter values that best fit the data. When the model is identified, then the fitting procedures used by typical software are a convenient way to estimate the model parameters. Additionally, some of these methods, e.g. full information maximum likelihood (FIML), have been shown to be at least as asymptotically efficient as any other consistent estimator. In other words, as the sample size goes to infinity, the FIML estimator has smaller than or equal variance to any other estimator.

However, if the model is not identified, then the program will typically be unable to fit the model and the procedure will exit with failure[7]. If they are unable to get the fitting procedures to converge without issue, researchers typically conclude that the model is not identifiable. While convenient, there are a number of disadvantages to determining identifiability of and estimating

---

[7]In other cases, a solution is generated, but it is inadmissible. When this happens, the program generally issues a warning (e.g. there is a Heywood case). In some cases, however, there is neither a warning in the output nor a Heywood case in the solution.
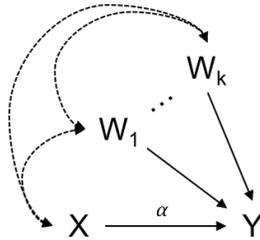
Figure 6: Path model where U_Y is independent of X given $W_1, ..., W_k$

model parameters in this way. First, if the procedure fails, it is not clear whether it failed due to identification or other issues (Kenny and Milan, 2012). Second, in rare cases, the procedure will provide solutions, even when the model is not identified. In such cases, the solutions will be incorrect, and the modeler may have no indiciation that this is the case. Third, when the model is not identified, the program is not helpful in indicating which parameters are not identifiable (Kenny and Milan, 2012). Fourth, since these methods attempt to simultaneously fit all of the model's parameters, misspecification may allow errors to propagate and bias parameter estimates throughout the model, rather than isolating the effects of misspecification locally (Bollen et al., 2007). Lastly, if the model is not identifiable, the researcher is only given indication of this fact after she has spent the effort and time to collect data (Kenny and Milan, 2012).

In this section, we introduce two basic graphical criteria for determining whether a given parameter is identified. When satisfied, these criteria also provide an expression for the parameter in terms of the covariance matrix, allowing the researcher to obtain an estimate for the parameter from the sample covariance matrix. As a result, even if the model is not identifiable, the researcher can still obtain estimates of parameters that are identifiable, and she knows which parameters are responsible for non-identifiability. Additionally, the estimates that are obtained using these methods are less likely to be biased by misspecification in unrelated parts of the model (Bollen et al., 2007). Lastly, graphical criteria for identification do not require data, so the researcher can determine the identifiability of the model or key parameters prior to collecting data.

3.1. Selecting Regressors

Consider the structural equation,

$$Y = \alpha X + \gamma_1 W_2 + ... + \gamma_k W_k + U_Y$$

and its corresponding path diagram shown in Figure 6. Because X is uncorrelated with U_Y given $W = \{W_1, ..., W_k\}$, we know that α is identified and equal to the coefficient of X in the regression of Y on $X, W_1, ..., W_k$. In other words, we can identify and estimate α by "controlling" or "adjusting" for the confounders, $W_1, ..., W_k$. This fact can also be seen from a graphical modeling perspective. By conditioning on W, we block all routes between X and Y except the path from X to Y. As a result, the partial correlation between X and Y given W is due only to the path $X \to Y$, and the regression coefficient, $\beta(Y, X|W)$ is equal to the path coefficient α.

It turns out that any set of variables, not just variables in the structural equation for Y, can potentially be used for control or adjustment in this manner. In general, if there exists a set of variables W that blocks all routes between X and Y other than the path $X \to Y$ and does not contain a descendant of Y, then $\alpha = \beta(Y, X|W)$ and can be estimated using ordinary least squares (OLS). This criterion, called single-door, characterizes when a path coefficient can be estimated using OLS[8].

Theorem 2. (Pearl, 2009, ch. 5.3.1) (Single-door Criterion) Let G be any acyclic causal graph in which α is the coefficient associated with the path $X \to Y$, and let G_α denote the diagram that results when $X \to Y$ is deleted from G. The coefficient α is identifiable if there exists a set of variables W such that (i) W contains no descendant of Y and (ii) W d-separates X from Y in G_α. If W satisfies these two conditions, then α is equal to the regression coefficient $\beta(Y, X|W)$[9], and we say that W is a *single-door admissible*. Conversely, if W does not satisfy these conditions, then $\beta(Y, X|W)$ is not equal to α (except in rare instances where spurious routes cancel).

In Figure 3a, we see that W blocks the back-door route $X \leftarrow Z \to W \to Y$ and X is d-separated from Y by W in Figure 3b. Therefore, α is identified and equal to $\beta(Y, X|W)$. This is to be expected since X is independent of U_Y given W in the structural equation, $Y = \alpha X + cW + U_Y$. (This independence can, of course, be verified by explicitly adding U_Y to the path diagram and applying d-separation, as in Figures 7a and 7b.) Theorem 2 tells us, however, that Z can also be used for adjustment since Z also d-separates X from Y in Figure 3b, and we obtain $\alpha = \beta(Y, X|Z)$[10]. Consider, however, Figure 3c. Variable Z satisfies the single-door criterion but W does not. Being a collider, W opens the route, $X \leftarrow Z \to W \leftrightarrow Y$, in violation of Theorem 2, leading to bias when adjusting for W. In conclusion, α is equal to $\beta(Y, X|Z)$ in Figures 3a and 3c. However, α is equal to $\beta(Y, X|W)$ in Figure 3a only.

The intuition for the requirement that Z not be a descendant of Y is depicted in Figures 7a and 7b. We typically do not display the error terms, which can be understood as latent causes. In Figure 7b, we show the error terms explicitly. It should now be clear that Y is a collider and conditioning on Z will create spurious correlation between X, U_Y, and Y leading to bias if adjusted for. This means that α can be estimated by the regression slope of Y on X, but adding Z to the regression equation would distort this slope, and yield a biased result.

---

[8]A variation of the single-door criterion, called the back-door criterion, graphically characterizes when a total effect can be identified and estimated using OLS (Pearl, 2009; Chen and Pearl, 2014). This criterion can be easily verified by inspecting the diagram and regarded as both a simplification and a generalization of the Alwin and Hauser (1975) method for the case where some errors are correlated.

[9]The estimand $\beta(Y, X|W)$ applies whether the variables are standardized or not.

[10]It turns out, however, that the choice of W is superior to that of Z in terms of estimation power (Kuroki and Miyakawa, 2003). The intuition here is that W is closer in the causal diagram (i.e. a more proximal cause) to Y, and, therefore, more effective at reducing variations in Y due to uncontrolled factors.
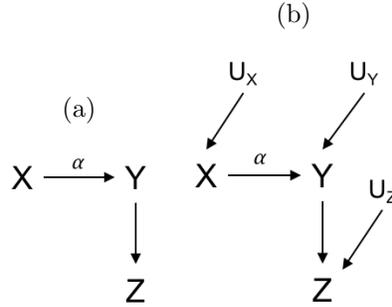
Figure 7: Example showing that adjusting for a descendant of Y induces bias in the estimation of α
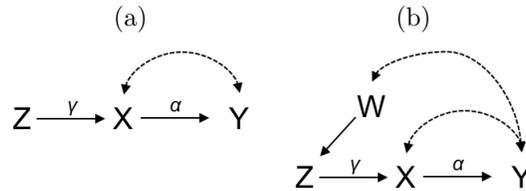
Figure 8: (a) Z qualifies as an instrumental variable (b) Z is an instrumental variable given W

3.2. Instrumental Variables

In Figure 8a, no single-door admissible set exists for α and it cannot be estimated using OLS. However, assuming without loss of generality that the variables are standardized and using Wright's equations, we see that $\sigma(Y, Z) = \gamma\alpha$ and $\sigma(X, Z) = \gamma$. As a result, $\alpha = \frac{\sigma(Y,Z)}{\sigma(X,Z)}$. In this case, we were able to identify α using Z as an *instrumental variable (IV)*. Typically, when α is identified using an IV, it is estimated using two-stage least squares (2SLS) regression.

In this example, we were able to identify α because the covariance between Z and Y could be factored into $\sigma(X, Z)$ and α, i.e. $\sigma(Y, Z) = \sigma(Y, X)\alpha$. Using Wright's rules, it is easy to see that this will be the case whenever each active route between Z and Y includes the path $X \to Y$. Thus, we can graphically characterize an IV for α, the coefficent from X to Y, as any variable Z that is d-separated from Y when we remove the path for α, $X \to Y$, from the diagram.

A natural question to ask is whether active routes between Z and Y that do not include $X \to Y$ can be blocked by conditioning. The answer is yes, so long as the conditioning set W does not contain any descendants of Y. The following is a graphical definition for instrumental variables that can be used to determine whether a given variable is an IV for a coefficient, conditional on W.

Definition 2. (Brito and Pearl, 2002) A variable Z is a *instrumental variable* given a set W for coefficient α (from X to Y) in a diagram G if

(i) W contains only non-descendants of Y

(ii) W d-separates Z from Y in the subdiagram G_α obtained by removing path $X \to Y$ from G

(iii) W does not d-separate Z from X in G_α

To demonstrate the usage of Definition 2, consider the following adapation of an example by Imbens (2014). Suppose that our goal is to estimate the effect α of a training program X on earnings Y. Since employees are not randomly assigned to the training program, there are a number of possible unobservable confounding factors that prevent us from directly estimating the causal effect. Employees that choose to participate may be more ambitious than employees that do not. Alternatively, employees that choose to participate may do so because they are suffering from poor performance.

Proximity to the training program Z is a possible instrumental variable since employees that live close to the training program are more likely to attend it. In this example, we will evaluate whether Z is indeed an instrument for α in four possible scenarios.

(i) The training program is located in the workplace. In this case, proximity Z may affect the numbers of hours employees spend at the office W since they spend less time commuting, and this, in turn, may affect their earnings Y.

(ii) Not only is the training program located in the workplace, but the efficiency of the workers (unmeasured) affects both the number of hours W and their salary Y.

(iii) While proximity may affect the number hours spent in the office, the number of hours spent at the office W does not affect earnings Y because workers in this industry can easily work from home. However, efficiency continues to affect both number of hours in the office W and salary Y.

(iv) Number of hours W again does not affect earning Y. However, work efficiency not only affects salary, but also whether they attend the program because less efficient workers are more likely to benefit from training.

Figures 9a-9d are the path diagrams for scenarios (i)-(iv) respectively. In each of these diagrams W is not a descendant of Y, and W does not separate Z from X in G_α. As a result, we only need to evaluate condition (ii) of Definition 2. In Figure 9a, we see that Z is a conditional instrument for α given W because Z is d-separated from Y given W in G_α. However, in Figure 9b, α cannot be identified using instruments. Attempting to block the spurious path $Z \to W \to Y$ activates another spurious path $Z \leftrightarrow W \leftrightarrow Y$. In Figure 9c, Z is an instrument for α as long as we do not condition on W since W is a collider on the path $Z \leftrightarrow W \leftrightarrow Y$. Similarly, in Figure 9d, Z is an instrument for α if we do not condition on W.

In the SEM literature, instrumental variables are typically defined relative to an equation rather than to a specific path or parameter as we have. By defining an instrument relative to a parameter, researchers can use instruments to identify and estimate individual parameters, even when the equation as a whole is not identified. Also, defining instruments relative to parameters, rather than equations, refines the conditions for when the equation, as a whole, is identifiable. Bollen (2012) notes that a necessary condition for an equation to be identified using instrumental
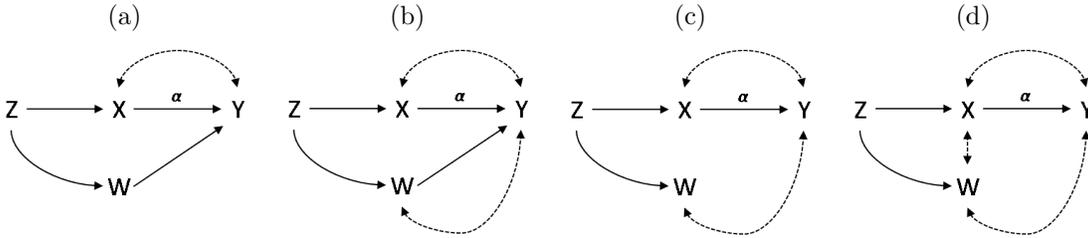
Figure 9: (a) Z is an conditional instrument for α given W (b) α is not identifiable using instruments (c) Z is an instrument for α (d) Z is an instrument for α

variables is that there are at least as many instruments as there are path coefficients in the structural equation; however, this condition is not sufficient. Instead, a sufficient condition is that each path coefficient in the equation has an instrumental variable. Further, a necessary and sufficient condition for the equation to be identified using IVs is that each path coefficient has an independent instrument. See Brito and Pearl (2002) definition for a precise, graph-based definition of independent instruments.

Lastly, we note that the single-door criterion and the graphical definition of instrumental variables can also be applied to SEMs with latent variables like CFA models. See Chen and Pearl (2014) for details.

3.3. Related Work

In theory, the identifiability of any linear-normal path model can be determined by analyzing the equations that characterize the implied covariance matrix in terms of the model parameters. If the model is partially recursive, these equations can be obtained using Wright's rules. If the model contains cycles, then they can also be obtained using matrix algebra (Bollen, 1989). Once these equations are obtained, determining whether any particular parameter can be solved for can be done using Gröbner basis methods, mathematical techniques for solving systems of polynomial equations. However, these methods are incredibly slow (a four-variable model can take over three days to solve using standard computers), and no efficient algorithm for solving this system of equations exists. Finding efficient algorithms for determining identifiability of linear-normal models is an area of ongoing research. Currently, there are a number of sufficient methods for identifying path models and their coefficients. However, none of them are complete. In other words, if they are unable to identify a particular coefficient, it may still be the case that the coefficient is identifiable, but the algorithm was not able to do so. In this subsection, we survey a number of recent graphical methods for identification, whose details are beyond the scope of this paper. Instead, we refer the reader to the relevant literature for more information.

In the section on Instrumental Variables, we mentioned that Brito and Pearl (2002) characterize when an equation can be identified using multiple instruments. Brito and Pearl (2006) generalized this *instrumental set* method to systematically identify all of the coefficients in a partially recursive model. A similar, but more powerful method was proposed by Foygel et al.

(2012) to determine identifiability of a larger set of models[11]. Additionally their criterion, called half-trek, applies to both partially recursive and non-recursive diagrams. The half-trek algorithm was further generalized by Chen et al. (2014), Drton and Weihs (2016), and Chen (2016). Lastly, Chen et al. (2016) and Chen et al. (2017) provided a method for generating new instrumental variables when some of the models parameters are known or identified. They used this method to provide an identification algorithm that is able to identify strictly more coefficients and models than the half-trek methods.

Other graphical methods for identification in linear path models include Tian (2005), Tian (2007), and Tian (2009). These methods identify parameters by converting the structural equations into orthogonal partial regression equations. Tian (2005) also gave a method for decomposing a complex path diagram into simpler sub-diagrams, showing that the original diagram is identifiable if and only if the sub-diagrams are.

Do-calculus (Pearl, 2009) and non-parametric algorithms for identifying causal effects (Tian and Pearl, 2002; Tian, 2002; Shpitser and Pearl, 2006; Huang and Valtorta, 2006), which have been proven to be complete for identifying causal effects in non-parametric models, may also be applied to parameter identification in path models (Chen and Pearl, 2014) (even when these path models do not represent causal relations).

Lastly, Rigdon (1995) described a set of necessary and sufficient graphical rules for the identification of nonrecursive models where the endogenous variables can be partitioned into sets of recursively-related blocks of size two or one.

4. Deriving Testable Implications

All methods for evaluating goodness of fit compare statistical constraints that are implied by the model specification with an assessment of the degree to which those constraints are satisfied in the data. The most common methods use computer software to first estimate the values of the model parameters from sample data and then, taking those estimates at face value, compute the model-implied covariance matrix. This model-implied covariance matrix is then compared to the sample covariance matrix using statistical tests, such as the well-known model chi-square test in SEM. The advantage of these methods are that they conveniently and simultaneously test all of the restrictions implied by the model on the covariance matrix.

However, there are a number of disadvantages, as well. First, they cannot be used unless the model is identified. Second, poor overall fit of the model does not provide the modeler with information about which aspect of the model needs to be revised[12]. Finally, it is possible that a global and simultaneous test will not reject the model, even when a crucial testable implication is violated. Global fit statistics, such as the model chi-square, Bentler Comparative Fit Index (CFI),

---

[11]Foygel et al. (2012) also released an R package implementing their algorithm called SEMID, which determines whether the entire model is identifiable given its path diagram.

[12] While modification indices can be used, they are restricted by the requirement that each modification results in an identified model. Additionally, they assume that, other than the specific path being investigated, the original model is correct. Lastly, they can be misleading if the model fails to fit for a reason other than a missing or wrong path, such as non-normality (Thoemmes et al., 2018).

and Steiger-Lind root mean square error of approximation (RMSEA), among others, measure only average or overall fit of the model to the sample covariance matrix. Thus, it can and does happen that values of global fit statistics mask violations of individual testable implications (Tomarken and Waller, 2003). Moreover, these masked violations may indicate a misspecification that is crucial for decision making.

In this section, we describe how individual testable implications of the model can be obtained from the graphical tools given in previous sections. These implications can be tested even when the model is not identified, the statistical power of each test is greater than that of a global test (McDonald, 2002; Bollen and Pearl, 2013), and, in the case of failure, the researcher knows exactly which implication of the model was violated, obtaining guidance when re-specifying the model.

When the model is identified, the results of these local tests should be reported in addition to the global fit statistics[13]. When the model is not identified, researchers can still evaluate its local fitness requirements and utilize the aforementioned methods of identifying and estimating individual parameters, which may be of significant research interest.

4.1. Vanishing Correlation Constraints

The method of d-separation allows modelers to derive conditional independence constraints implied by the path model simply by inspecting the diagram. In the case of recursive models without equality constraints, the number of implied conditional independences exactly equals the model degrees of freedom, or df_M. For example, if $df_M = 5$ for a recursive model, then there are a total of five partial correlations that should vanish (equal zero), given the configuration of paths specified in the model. These five vanishing partial correlations represent all the constraints that the model specification imposes on the covariance matrix. As a result, they represent all the testable implications of the model, assuming that the endogenous variables are normally distributed (Geiger and Pearl, 1993). The model chi-square test, where $df_M = 5$ for the same graph, can be seen as an overall significance test of whether all five partial correlations vanish, but inspecting the values of the individual sample partial correlations provides all the details behind the global chi-square test. Additionally, since these conditional independence constraints are derived, using d-separation, from specific aspects of the model specification, the modeler knows which aspects of the model are incorrect when one of these constraints is violated. It is in this sense that we say such tests assess "local" rather than "global" fit.

For example, in Figure 10a, we obtain the following vanishing partial correlations: $\rho(V_2, V_3 | V_1) = 0$, $\rho(V_1, V_4 | V_2, V_3) = 0$, $\rho(V_2, V_5 | V_4) = 0$, and $\rho(V_3, V_5 | V_4) = 0$. Hypothesis testing whether these partial correlations are equal to zero from data can be accomplished using the Fisher Z transformation, which is implemented in many standard statistical computer tools and the DAGitty R package (Textor et al., 2011; Thoemmes et al., 2018). If a test shows that a constraint, say $\rho_{V_2 V_3 \cdot V_1} = 0$ does not hold in the data, we have reason to believe that the model specification is incorrect and should reconsider the lack of a path between V_2 and V_3.

In some cases, vanishing partial correlations derived from d-separation may be redundant. In

---

[13]A serious flaw in many, if not most, published SEM studies is the failure to report information about local fit (Vermeulen and Hartmann, 2015; Goodboy and Kline, 2017).
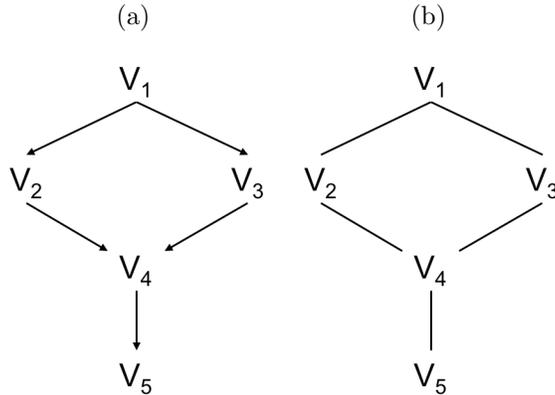
Figure 10: (a) Example illustrating vanishing partial correlation (b) The skeleton of the model in (a)

other words, some may be implied by others. There are ways to derive minimal sets of implied vanishing partial correlations, called basis sets, that consist of the smallest number of conditional independences that imply all others for a particular diagram. For the same diagram, there may be multiple basis sets, each with the same overall number of conditional independences, that can be derived using different but logically-consistent methods to generate a basis set. We do not cover basis sets here, but see Kang and Tian (2009), Pearl (2009, pp. 142-145), or Shipley (2000, pp. 61-63) for more information.

Lastly, d-separation can also be used to derive conditional independence constraints in latent variable models. See Thoemmes et al. (2018) for details, as well as additional graphical tools for local fit evaluation of latent variable models.

4.2. Overidentifying Constraints

In the previous subsection, it was noted that all of the testable implications of a recursive and normal path model take the form of vanishing partial correlation constraints that can be derived using d-separation. Non-recursive and partially recursive models, however, may imply other types of constraints on the covariance matrix. Some of these constraints can be derived by identifying parameters using two or more distinct identification strategies[14]. Such parameters are often called *overidentified*.

Consider the path diagram shown in Figure 11. Since W satisfies the single-door criterion for α, the path model implies that $\alpha = \beta(Y, X|W)$. Additionally, Z is an IV for α so the model also implies that $\alpha = \frac{\beta(Y,Z)}{\beta(X,Z)}$. In this case, we say that α is *overidentified*, and we obtain a testable implication of the model by equating the two expressions, $\beta(Y, X|W)$ and $\frac{\beta(Y,Z)}{\beta(X,Z)}$. If the data does not corroborate this equality, then the researcher has reason to believe that the model is misspecified, and it does not fit the data. Specifically, there is evidence that the model restrictions

---

[14]See Pearl (2004) for a formal definition of "distinct identification strategies".
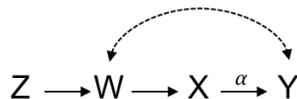
Figure 11: α is overidentified using W as a single-door admissible set and Z as an IV

that allow Z to be an IV (e.g. the lack of a path between Z and Y) are incorrect, or the model restrictions that allow $\{W\}$ to be a single-door admissible set (e.g. the lack of a bidirected path between W and X and the lack of a bidirected path between X and Y) are, or both.

The single-door criterion and IVs provide researchers with a systematic way to derive these overidentifying constraints. (See Chen et al. (2017) for graphical algorithm that discovers overidentifying constraints using more complex identification methods). Hypothesis testing can then be conducted using the Durbin-Wu-Hausman test, also called the Hausman specification test (Hausman, 1978). This test is commonly implemented in standard statistical computer tools.

When two single-door admissible sets exist for a coefficient, the corresponding overidentifying constraint will always be equivalent to a vanishing partial correlation constraint that could also have been derived using d-separation (Chen and Pearl, 2015). In contrast, overidentifying constraints derived from two different IVs or a single-door admissible set and an IV, may yield other types of constraints that cannot be obtained from d-separation.

4.3. Equivalent Models

Since conditional independences represent all of the constraints a recursive path model imposes on the covariance matrix, two recursive models that share the same implied conditional independences cannot be distinguished using the sample covariance matrix. In such cases, we say that the models are covariance equivalent. If the endogenous variables are normally distributed, then covariance equivalence implies general equivalence and the models cannot be distinguished from data at all. Most authors of SEM studies fail to acknowledge the existence of equivalent models, which is a serious form of confirmation bias (Hoyle and Isherwood, 2013).

Researchers can check whether two models share the same implied conditional independences using d-separation. However, in complex models with many variables, applying the d-separation criterion by hand can be tedious. Instead, if the diagram does not contain bidirected paths, model equivalence can be checked by simply noting whether the two models have the same *skeleton* and *v-structures* (Pearl, 2009, pp.145-149). (If the diagrams have bidirected paths, we can still apply this method by first replacing each bidirected path with a latent common cause. For example, the bidirected path $X \leftarrow\!\dashrightarrow Y$ is replaced with $X \leftarrow U \rightarrow Y$.) The skeleton is the undirected diagram obtained by replacing all arrows with undirected paths, which are depicted in the diagram using $--$. For example, the skeleton for Figure 10a is Figure 10b. V-structures are two converging arrows whose tails are not connected by an arrow. For example, $V_2 \rightarrow V_4 \leftarrow V_3$ is the only v-structure in Figure 10a.

Theorem 3. (Verma and Pearl, 1990) Two recursive and normal path models are covariance equivalent if and only if they entail the same sets of zero partial correlations. Moreover, two such
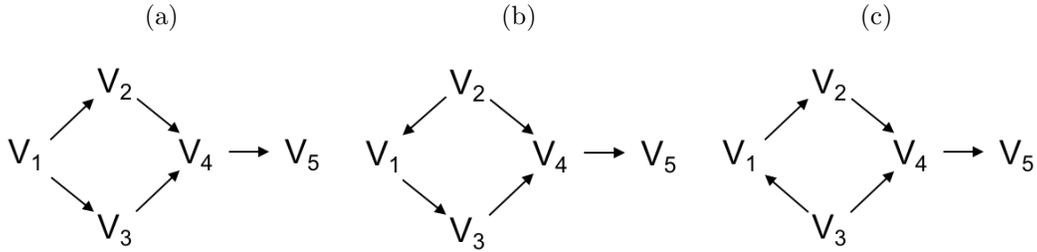
Figure 12: Models (a), (b), and (c) are equivalent.
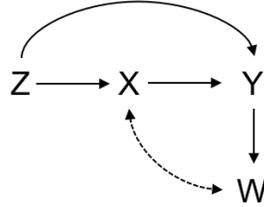
Figure 13: Counterexample to the standard Replacement Rule; The arrow $X \to Y$ cannot be replaced.

models are covariance equivalent if and only if their corresponding graphs have the same skeletons and the same sets of v-structures.

The graphs in Figures 12a, 12b, and 12c are equivalent because they share the same skeleton and v-structures. Note that we cannot reverse the path from V_4 to V_5 since doing so would generate a new v-structure, $V_2 \to V_4 \leftarrow V_5$.

The graphical criterion given in Theorem 3 is necessary and sufficient for equivalence between recursive path models with normally-distributed endogenous variables. It is a necessary condition for equivalence between non-recursive models since d-separation in the diagram implies vanishing partial correlation in the covariance matrix. In contrast, the more prevalent replacement criterion (Lee and Hershberger, 1990) is not always valid[15]. Pearl (2012) gave the following example depicted in Figure 13. According to the replacement criterion, we can replace the path $X \to Y$ with a bidirected path $X \leftrightarrow Y$ and obtain a covariance equivalent model when all predictors (Z) of the effect variable (Y) are the same as those for the source variable (X). Unfortunately, the post-replacement model imposes the constraint, $\rho(WZ|Y) = 0$, which is not imposed by the original model. This can be seen from the fact that, conditioned on Y, the route $Z \to Y \leftarrow X \leftrightarrow W$ is unblocked and becomes blocked if replaced by $Z \to Y \leftrightarrow X \leftrightarrow W$. The same applies to the route $Z \to X \leftrightarrow W$, since Y would cease to be a descendant of X.

---

[15]The replacement rule violates the transitivity of equivalence (Hershberger and Marcoulides, 2006), yet it is still used in most of the SEM literature (Mulaik, 2009; Williams, 2012, pp. 247-260).

5. Computer Tools

There are freely-available computer tools for analyzing directed acyclic graphs that automatically apply the graphical methods just described. Directed acyclic graphs (DAGs) corresponds to a recursive or partially recursive path diagram. In other words, they do not allow cycles. While the tools presented in this paper can be additionally applied to cyclic graphs and fully nonrecursive path models, much of the software developed has focused on directed acyclic graphs.

Some of these computer tools are stand-alone applications for installation on personal computers, but others are online applications that analyze graphs drawn by the user in an Internet browser. All such tools support the evaluation of a path diagram in the planning stages for a study. Some of these tools are described next. This list is not comprehensive, but all these tools help researchers to reap the potential benefits of analyzing their graphs and thus testing their ideas before collecting the data:

1. The DAGitty program is web-based tool for analyzing causal graphs that can also be used offline (Textor et al., 2011). It evaluates whether direct or total effects are identified either through covariate selection or through the analysis of instruments. It also lists conditional independences implied by the graph.

2. Textor et al. (2011) also developed a DAGitty R package that includes the functionality of the web-based tool. Thoemmes et al. (2018) extended this package to include tests for vanishing partial correlation and to implement additional graphical methods for analysis of latent variable models.

3. The Belief and Decision Network Tool (Knoll et al., 2008) is a Java applet for learning about the concept of d-separation. For example, after drawing a graph onscreen, this program can then be optionally run in ask the applet mode, where the user clicks on two focal variables and a set of covariates, and the program automatically indicates whether the focal variables are independent, given those covariates.

4. The dagR package for R (Breitling, 2010) provides a set of functions for drawing, manipulating, and analyzing directed acyclic graphs and also simulating data consistent with the corresponding diagram. It can evaluate effects of analyzing different subsets of covariates when estimating causal effects of exposure variables on outcome variables, among other capabilities. Graphs are specified in syntax, but the corresponding graph of the model can be manipulated in the R environment.

6. Example Problem

In this section, we present an illustrative example to demonstrate how typical analysis can be expanded and improved by incorporating ideas from graphical modeling, thereby creating an expanded and improved overall practice. Presented in Figure 14a is a variation on a path model analyzed by Roth et al. (1989). We assume a linear model with continuous variables. The hypotheses are that
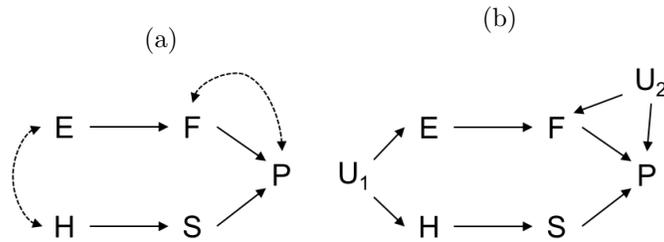
Figure 14: (a) Example path model, where E represents exercise, H mental hardiness, F physical fitness, and S stress (b) The same model as it would be expressed in the DAGitty computer tool, where U_1 and U_2 are latent variables

1. exercise (E) and mental hardiness (H) covary;

2. variables E and H each indirectly affects health problems (P) through, respectively, physical fitness (F) and stress (S); and

3. the errors of F and P covary; that is, they share an unmeasured common cause.

Standard practice dictates that once the model has been specified, the best fit for the model parameters should be determined using data. If the model is not identifiable, then this procedure would fail. At this point, researchers would be forced to respecify the model or collect data on additional variables that may enable identification. The former is clearly not an ideal solution because model specification should be based on reality and not used as a workaround for identification issues. The latter can be difficult to accomplish since standard SEM programs do not provide the researcher with pointers to the variables that would enable identification. Thus, the researcher may take the time and resources to collect data on new variables, only to find that the model is still not identified.

In contrast, graphical tools may enable researchers to estimate coefficients of interest (e.g. the effect of stress, exercise, or mental hardiness on health problems) even when the model, as a whole, is not identified. If these quantities are also not identifiable, then the researcher can still use graphical tools to determine which additional variables would enable identification. She would simply add the variable in question to the path diagram and use the tools described above to determine whether the model or coefficients of interest are now identified.

Next, we describe how the graphical identification methods described in this paper can be used to identify and estimate each of the coefficients in Figure 14a. Readers can verify the identification results and testable implications for this example using DAGitty. Error correlations are represented in DAGitty by specifying latent variables as common causes, such as U_1 and U_2 in Figure 14b. See the appendix for code that automatically generates the graph for this example in DAGitty.

The total number of available identification stategies reported by DAGitty is indicated in parentheses for each coefficient listed next:

1. $F \to P$ (3). The DAGitty tool indicates that the coefficient for this direct effect is identified through the instruments E given S and H given S. Controlling for S in each instrument

Coefficient	OLS		2SLS	
			Estimator	
$E \to F$	.108	$(\varnothing)$	-.646	(H)
			.719	(S)
$H \to S$	-.203	$(\varnothing)$	1.469	(E)
			-1.637	(F)
$F \to P$	—		-.558	$(E|S)$
			-6.927	$(H|S)$
			-.734	$(E|H)$
$S \to P$	.628	(E)	1.161	$(H|F,E)$
	.597	(H)		

Table 1: Estimators of coefficients for Figure 14

d-separates variables E and H from the outcome P in the modified graph without $F \to P$. Similarly, the direct effect is identified by the instrument E given H.

2. $S \to P$ (3). Two different single-door admissible sets, E and H, identify the direct effect: Controlling for either E or H blocks the back-door route $S \leftarrow H \leftarrow\!\dashrightarrow E \to F \to P$ in the modified graph without $S \to P$. The instrument H given $\{F, E\}$ also identifies the direct effect.

3. $E \to F$, $H \to S$ (3 each). There are no back-door treks between either pair of explanatory and outcome variables; thus, the coefficient for each direct effect can be estimated in bivariate regression (i.e., $\varnothing$, the empty set, is single-door admissible). Two different instruments are also available for each direct effect, H and S for $E \to F$ and E and F for $H \to S$.

A veritable cornucopia of identification strategies are available for this example by applying concepts from graphical modeling. Summary statistics (correlations, standard deviations, and means) for the observed variables in Figure 14a measured within a sample of 373 university students were analyzed in SPSS in order to generate results for this example. Syntax files with the data for this analysis summarized in matrix form and output files for SPSS in both plain text and Adobe PDF format can be downloaded from the supplemental materials page for this article.

The estimate for $X \to Y$ using Z as an instrument can be obtained by conducting 2SLS regression or directly from the summary statistics by dividing $\hat{\sigma}(Y, Z)$ by $\hat{\sigma}(X, Z)$, where $\hat{\sigma}(Y, Z)$ is the sample covariance between Y and Z and $\hat{\sigma}(X, Z)$ is the sample covariance between X and Z. Similarly,

$$\frac{\hat{\sigma}(Y, Z|W)}{\hat{\sigma}(X, Z|W)} \tag{12}$$

is the estimate for $X \to Y$ using an instrument, Z given W.

Reported in the top part of Table 1 are unstandardized estimates of direct effects for this example with conditioning sets given in parenthesis. Since the model has been shown to be identified, estimates could also be obtained using ML procedures implemented in standard SEM

software. As mentioned above, some of these procedures like FIML have been shown to be at least as asymptotically efficient as any other estimator and, therefore, may be preferable to the estimates provided in Table 1. However, researchers should verify, to the best of their abilities, that the model specification is correct before using estimates derived from graphical or ML methods.

Typically, researchers inspect the values of global fit statistics provided by standard SEM tools to assess the model's fit. Values of selected global fit statistics in ML estimation, obtained using LISREL and Mplus, for this example are listed next:

$$\chi^2_M(4) = 10.529, p = .032$$
$$\text{RMSEA} = .066, 90\% \text{ CI } [.017, .116]$$
$$\text{CFI} = .958; \text{SRMR} = .051$$

The model chi-square is marginally significant at the .05 level, and the upper bound of the confidence interval based on the RMSEA is unfavorable, but other results do not indicate grossly poor global fit. Through selective reporting (e.g., omit the CI based on the RMSEA) and ignoring the failed chi-square test–both common practices in the SEM literature (Ropovik, 2015)–a researcher could potentially argue for retaining the model.

Even if none of the statistics indicate poor fit, it is still important that researchers test the model using the graphical methods described in this paper. Global fit statistics may fail to detect misspecification that can greatly bias estimates of important coefficients because they are overall summaries of the model's fit (Tomarken and Waller, 2003). In contrast, graphical methods allow researchers to derive the testable implications that underlie global fit statistics and test them individually.

If the researcher does find the global fit statistics to be problematic, d-separation and the above identification criteria also enable the researcher to determine exactly which testable implications are violated and revise the model accordingly. In contrast, the modification indices provided by standard SEM software are restricted by the requirement that each modification results in an identified model. Additionally, they assume that, other than the specific path being investigated, the original model is correct. Lastly, they can be misleading if the model fails to fit for a reason other than a missing or wrong path, such as non-normality (Thoemmes et al., 2018). Next, we demonstrate how to use graphical tools to test and revise the model shown in Figure 14a.

Inspecting Table 1 immediately reveals issues with the model fit. Estimates for $E \rightarrow F$ and $H \rightarrow S$ vary appreciably in both direction and magnitude, and thus are inconsistent. This pattern attests to likely specification error. For example, the individual 2SLS estimates for $E \rightarrow F$, $-.646$ and $.719$ (see the table), each assumes that the corresponding instrument, respectively, H and S, has no direct effect on F. Additionally, they assume that there is no bidirectional route that connects either instrument just mentioned with F (see Figure 14a). At least one of these assumptions is wrong and the researcher can use this information to revise the model accordingly.

Estimates of $F \rightarrow P$ agree in direction (all negative) but appreciably vary in magnitude; the range is -6.927 to -.558 (see Table 1). The outlier value, or -6.927, is from the 2SLS analysis that assumes no direct effect between variables H and P. Values of multiple estimates for $S \rightarrow P$ are all > 0 and generally similar in magnitude (.597 to 1.161), which is more reassuring.

Conditional Independence	Partial Correlation	p
$E \perp S\|H$	-0.058	0.264
$H \perp F\|E$	0.089	0.086
$F \perp S\|H$	-0.117	0.024
$F \perp S\|E$	-0.120	0.020
$H \perp P\|E,S$	-0.093	0.074

Table 2: Implied conditional independences for Figure 14
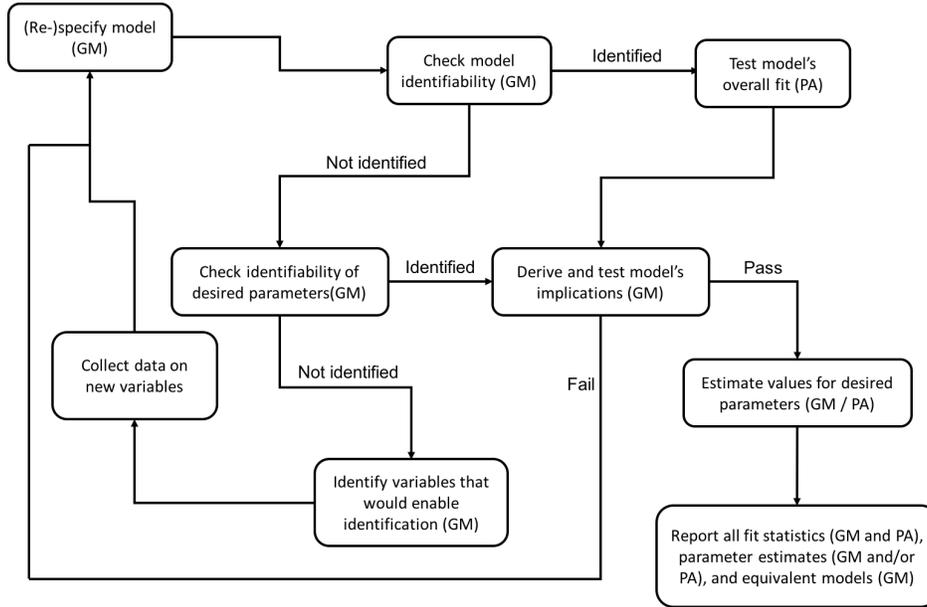
Figure 15: Flow chart depicting expanded practice that combines techniques from graphical modeling and standard path analysis. Steps utilizing graphical modeling and standard path analytic techniques indicated by (GM) and (PA) respectively.

Applying d-separation also reveals implications that can be use to test the model. Listed next are five vanishing partial correlations implied by Figure 14a:

$$\rho(E,S|H) = \rho(H,F|E) = \rho(F,S|E) = \rho(F,S|H) = \rho(H,P|ES) = 0$$

Table 2 shows the sample partial correlations and their p-value for the hypothesis that the population partial correlations are equal to zero. The tests for $\rho(F,S|E) = 0$ and $\rho(F,S|H) = 0$ fail at the 5% significance level, suggesting that F and S should not be d-separated by either E nor H. The simplest revision to the model specification would include an edge between F and S. The lack of this edge in Figure 14a could also explain the inconsistent 2SLS estimates for $E \rightarrow F$, $H \rightarrow S$, and $F \rightarrow P$ shown in Table 1.

Once the model has been revised, the researcher should repeat the procedure described above

until she obtains a model that passes both global and local tests for goodness of fit. At this point, it is important to consider equivalent models, which can be derived using the graphical procedures described above. If she still feels confident that the model specification is correct, then estimated values for the model parameters can be obtained using either standard SEM software or OLS/2SLS. The former may be preferable when the model specification is correct due to better efficiency. The latter may be preferable when the researcher is unsure about the model specification due to equivalent models or borderline results for fit tests, since they may isolate errors rather than allowing them to propagate (Bollen et al., 2007). This procedure is summarized in a flow chart depicted in Figure 15.

7. Conclusion

The benefits of causal graphs are usually attributed to their ability to represent theoretical assumptions visibly and transparently, by abstracting away unnecessary algebraic details. Not as widely recognized among practitioners of SEM is the inferential power of analyzing graphs apart from analyzing the data. We demonstrated how a few basic graphical tools can be applied to a wide variety of modeling tasks in experimental psychology. Specifically, we showed how these tools can determine the identification status of path coefficients through both partial regression and IV methods. We further demonstrated how to derive testable implications through overidentification and the reading of vanishing partial correlations, both giving rise to new methods for assessing local fit. Finally, we discussed the identification of equivalent models, and illustrated the application of these tools and methods through numerical examles.

Appendices

A. DAGitty Model Code for Figure 14a

Copy the text below and paste it in the model code window of DAGitty in order for the path diagram to be to automatically drawn. The blank line must be retained in order for the code to correctly execute.

```
E  1  @−1.300,−0.500
F  1  @−0.700,−0.500
H  1  @−1.300,0.300
P  1  @−0.200,−0.100
S  1  @−0.700,0.300
U1 U  @−1.800,−0.100
U2 U  @−0.250,−0.450

E  F
F  P
H  S
S  P
```

U1 E H
U2 F P

References

ALWIN, D. and HAUSER, R. (1975). The decomposition of effects in path analysis. *American Sociological Review* **40** 37–47.

BLALOCK, H. (1962). Four-variable causal models and partial correlations. *American Journal of Sociology* **68** 182–194.

BOLLEN, K. (1989). *Structural Equations with Latent Variables.* John Wiley, New York.

BOLLEN, K. and PEARL, J. (2013). Eight myths about causality and structural equation models. In *Handbook of Causal Analysis for Social Research* (S. Morgan, ed.). Springer, 301–328.

BOLLEN, K. A. (2012). Instrumental variables in sociology and the social sciences. *Annual Review of Sociology* **38** 37–72.

BOLLEN, K. A., KIRBY, J. B., CURRAN, P. J., PAXTON, P. M. and CHEN, F. (2007). Latent variable models under misspecification: two-stage least squares (2sls) and maximum likelihood (ml) estimators. *Sociological Methods & Research* **36** 48–86.

BREITLING, L. P. (2010). dagr: a suite of r functions for directed acyclic graphs. *Epidemiology* **21** 586–587.

BRITO, C. and PEARL, J. (2002). Generalized instrumental variables. In *Uncertainty in Artificial Intelligence, Proceedings of the Eighteenth Conference* (A. Darwiche and N. Friedman, eds.). Morgan Kaufmann, San Francisco, 85–93.

BRITO, C. and PEARL, J. (2006). Graphical condition for identification in recursive SEM. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence.* AUAI Press, Corvallis, OR, 47–54.

CHEN, B. (2016). Identification and overidentification of linear structural equation models. In *Advances in Neural Information Processing Systems.*

CHEN, B., BAREINBOIM, E. and J., P. (2016). Incorporating knowledge into structural equation models using auxiliary variables. In *Proceedings of the Twenty-fifth International Joint Conference on Artificial Intelligence* (S. Kambhampati, ed.). AAAI Press, Palo Alto, CA, 3577–3583.

CHEN, B., KUMOR, D. and BAREINBOIM, E. (2017). Identification and model testing in linear structural equation models using auxiliary variables. In *Proceedings of the 34th International Conference on Machine Learning* (D. Precup and Y. W. Teh, eds.), vol. 70 of *Proceedings of Machine Learning Research.* PMLR, International Convention Centre, Sydney, Australia. URL `http://proceedings.mlr.press/v70/chen17f.html`

CHEN, B. and PEARL, J. (2014). Graphical tools for linear structural equation modeling. Tech. Rep. R-432, <http://ftp.cs.ucla.edu/pub/stat_ser/r432.pdf>, Department of Computer Science, University of California, Los Angeles, CA.

CHEN, B. and PEARL, J. (2015). Exogeneity and robustness. Tech. Rep. R-449, <http://ftp.cs.ucla.edu/pub/stat_ser/r449.pdf>, Department of Computer Science, University of California, Los Angeles, CA. Submitted.

CHEN, B., TIAN, J. and PEARL, J. (2014). Testable implications of linear structual equation models. In *Proceedings of the Twenty-eighth AAAI Conference on Artificial Intelligence* (C. E. Brodley and P. Stone, eds.). AAAI Press, Palo, CA. <http://ftp.cs.ucla.edu/pub/stat_ser/r428-reprint.pdf>.

CRÁMER, H. (1946). *Mathematical Methods of Statistics.* Princeton University Press, Princeton, NJ.

DRTON, M. and WEIHS, L. (2016). Generic identifiability of linear structural equation models by ancestor decomposition. *Scandinavian Journal of Statistics* **43** 1035–1045.

ELWERT, F. (2013). Graphical causal models. In *Handbook of causal analysis for social research.* Springer, 245–273.

FOYGEL, R., DRAISMA, J. and DRTON, M. (2012). Half-trek criterion for generic identifiability of linear structural equation models. *The Annals of Statistics* **40** 1682–1713.

GEIGER, D. and PEARL, J. (1993). Logical and algorithmic properties of conditional independence and graphical models. *The Annals of Statistics* 2001–2021.

GOODBOY, A. K. and KLINE, R. B. (2017). Statistical and practical concerns with published communication research featuring structural equation modeling. *Communication Research Reports* **34** 68–77.

HAUSMAN, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society* 1251–1271.

HAYDUK, L., CUMMINGS, G., STRATKOTTER, R., NIMMO, M., GRYGORYEV, K., DOSMAN, D., GILLESPIE, M., PAZDERKA-ROBINSON, H. and BOADU, K. (2003). Pearl's d-separation: One more step into causal thinking. *Structural Equation Modeling* **10** 289–311.

HERSHBERGER, S. L. and MARCOULIDES, G. A. (2006). The problem of equivalent structural models. In *Structural equation modeling: A second course* (G. R. Hancock and R. O. Mueller, eds.). Information Age Publishing, Greenwich, CT, 3–37.

HOYLE, R. H. (2012). *Handbook of structural equation modeling.* Guilford Press.

HOYLE, R. H. and ISHERWOOD, J. C. (2013). Reporting results from structural equation modeling analyses in archives of scientific psychology. *Archives of scientific psychology* **1** 14.

HUANG, Y. and VALTORTA, M. (2006). Pearl's calculus of intervention is complete. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence* (R. Dechter and T. Richardson, eds.). AUAI Press, Corvallis, OR, 217–224.

IMBENS, G. (2014). Instrumental variables: An econometrician's perspective. Tech. rep., National Bureau of Economic Research.

KANG, C. and TIAN, J. (2009). Markov properties for linear causal models with correlated errors. *The Journal of Machine Learning Research* **10** 41–70.

KENNY, D. A. and MILAN, S. (2012). Identification: A non-technical discussion of a technical issue. In *Handbook of Structural Equation Modeling* (G. M. S. W. R. Hoyle, D. Kaplan, ed.). Guilford Press, New York, 145–163.

KLINE, R. (2016). *Principles and Practice of Structural Equation Modeling.* 4th ed. Methodology in the social sciences.
URL http://books.google.com/books?id=-MDPILyu3DAC

KNOLL, B., KISYNSKI, J., CARENINI, G., CONATI, C., MACKWORTH, A. and POOLE, D. (2008). Aispace: Interactive tools for learning artificial intelligence. In *Proc. AAAI 2008 AI Education Workshop.*

KOLLER, D. and FRIEDMAN, N. (2009). *Probabilistic graphical models: principles and techniques.* MIT press.

KUROKI, M. and MIYAKAWA, M. (2003). Covariate selection for estimating the causal effect of control plans by using causal diagrams. *Journal of the Japanese Royal Statistical Society, Series B* **65** 209–222.

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

MCDONALD, R. (2002). What can we learn from the path equations?: Identifiability constraints, equivalence. *Psychometrika* **67** 225–249.

MORGAN, S. and WINSHIP, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research (Analytical Methods for Social Research).* Cambridge University Press, New York, NY.

MULAIK, S. A. (2009). *Linear causal modeling with structural equations.* CRC Press.

PEARL, J. (1995). Causal diagrams for empirical research. *Biometrika* **82** 669–710.

PEARL, J. (1998). Graphs, causality, and structural equation models. *Sociological Methods and Research* **27** 226–284.

PEARL, J. (2004). Robustness of causal claims. In *Proceedings of the 20th conference on Uncertainty in artificial intelligence.* AUAI Press.

PEARL, J. (2009). *Causality: Models, Reasoning, and Inference.* 2nd ed. Cambridge University Press, New York.

PEARL, J. (2012). The causal foundations of structural equation modeling. In *Handbook of Structural Equation Modeling* (R. Hoyle, ed.). Guilford Press, New York, 68–91.

PEARL, J. (2014). The deductive approach to causal inference. *Journal of Causal Inference* **2** 115–129.

RIGDON, E. E. (1995). A necessary and sufficient identification rule for structural models estimated in practice. *Multivariate Behavioral Research* **30** 359–383.

ROPOVIK, I. (2015). A cautionary note on testing latent variable models. *Frontiers in psychology* **6**.

ROTH, D. L., WIEBE, D. J., FILLINGIM, R. B. and SHAY, K. A. (1989). Life events, fitness, hardiness, and health: A simultaneous analysis of proposed stress-resistance effects. *Journal of Personality and Social Psychology* **57** 136.

SHIPLEY, B. (2000). A new inferential test for path models based on directed acyclic graphs. *Structural Equation Modeling* **7** 206–218.

SHPITSER, I. and PEARL, J. (2006). Identification of conditional interventional distributions. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence* (R. Dechter and T. Richardson, eds.). AUAI Press, Corvallis, OR, 437–444.

SPIRTES, P. (1995). Directed cyclic graphical representation of feedback. In *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence* (P.Besnard and S. Hanks, eds.). Morgan Kaufmann, San Mateo, CA, 491–498.

TEXTOR, J., HARDT, J. and KNÜPPEL, S. (2011). Dagitty: a graphical tool for analyzing causal diagrams. *Epidemiology* **22** 745.

THOEMMES, F., ROSSEEL, Y. and TEXTOR, J. (2018). Local fit evaluation of structural equation models using graphical criteria. *Psychological methods* **23** 27.

TIAN, J. (2002). *Studies in Causal Reasoning and Learning.* Ph.D. thesis, Computer Science Department, University of California, Los Angeles, CA.

TIAN, J. (2005). Identifying direct causal effects in linear models. In *Proceedings of the National Conference on Artificial Intelligence*, vol. 20. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999.

TIAN, J. (2007). A criterion for parameter identification in structural equation models. In *Proceedings of the Twenty-Third Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-07)*. AUAI Press, Corvallis, Oregon.

TIAN, J. (2009). Parameter identification in a class of linear structural equation models. In *Proceedings of the Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09)*.

TIAN, J. and PEARL, J. (2002). A general identification condition for causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*. AAAI Press/The MIT Press, Menlo Park, CA, 567–573.

TOMARKEN, A. J. and WALLER, N. G. (2003). Potential problems with "well fitting" models. *Journal of Abnormal Psychology* **112** 578.

VANDERWEELE, T. (2015). *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press.

VERMA, T. and PEARL, J. (1990). Equivalence and synthesis of causal models. In *Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence*. Cambridge, MA. Also in P. Bonissone, M. Henrion, L.N. Kanal and J.F. Lemmer (Eds.), *Uncertainty in Artificial Intelligence 6*, Elsevier Science Publishers, B.V., 255–268, 1991.

VERMEULEN, I. and HARTMANN, T. (2015). Questionable research and publication practices in communication science. *Communication Method and Measures* **9** 189–192.

WILLIAMS, L. (2012). Equivalent models: Concepts, problems, alternatives. *Handbook of Structural Equation Modeling (R. Hoyle, ed.). Guilford Press, New York* 247–260.

WOLFLE, L. M. (1980). Strategies of path analysis. *American Educational Research Journal* **17** 183–209.

WRIGHT, S. (1921). Correlation and causation. *Journal of Agricultural Research* **20** 557–585.