Using Descendants as Instrumental Variables for the Identification of Direct Causal Effects in Linear SEMs

Hei Chan
Institute of Statistical Mathematics,
10-3 Midori-cho, Tachikawa,
Tokyo 190-8562, Japan
chanheigeorge@gmail.com

Manabu Kuroki
Department of Systems Innovation,
Graduate School of Engineering Science,
Osaka University,
1-3 Machikaneyama-cho, Toyonaka,
Osaka 560-8531, Japan
mkuroki@sigmath.es.osaka-u.ac.jp

Abstract

In this paper, we present an extended set of graphical criteria for the identification of direct causal effects in linear Structural Equation Models (SEMs). Previous methods of graphical identification of direct causal effects in linear SEMs include methods such as the single-door criterion, the instrumental variable and the IV-pair, and the accessory set. However, there remain graphical models where a direct causal effect can be identified and these graphical criteria all fail. As a result, we introduce a new set of graphical criteria which uses descendants of either the cause variable or the effect variable as "path-specific instrumental variables" for the identification of the direct causal effect as long as certain conditions are satisfied. These conditions are based on edge removal and the existing graphical criteria of instrumental variables, and the identifiability of certain other total effects, and thus can be easily checked.

1 Introduction

Structural Equation Models (SEM) is a useful tool for causal analysis, and is widely used in areas of social science such as economics (Bollen 1989; Duncan 1975). Research by scientists, social scientists, and computer scientists in this area has allowed the problem to be applied in real-life models.

Appearing in Proceedings of the 13^{th} International Conference on Artificial Intelligence and Statistics (AISTATS) 2010, Chia Laguna Resort, Sardinia, Italy. Volume 9 of JMLR: W&CP 9.

In a linear SEM, the relationships between observed variables are expressed in linear equations. The structure of the equations is such that they not only express the linear relationships between the variables, together with a stochastic error term for unobserved factors, but also the causal dependence among the observed variables. For each variable Y, its structural equation where it appears on the left-hand side, the presence (and absence) of a variable X on the right-hand side specifies that X is (or is not) a direct cause of Y.

A fundamental problem in linear SEMs is to estimate the strength of a certain direct causal effect from one variable to another from a combination of observed data and model structure. This is called the *identification problem* (Fisher 1966). Although many methods, both algebraic and graphical, have been developed over the years, this problem is still not solved. Currently there are no sufficient or necessary criteria for deciding whether a causal effect can be identified from observed data. Current identification methods based on graphical criteria, including the single-door criterion, the front door criterion and the back door criterion, the instrumental variable and the IV-pair, and the accessory set, can be used to identify direct causal effects only when certain conditions are met. However, there exist direct causal effects that can be found using algebraic methods by solving for a set of equations involving the causal effects and the covariances but cannot be identified using these methods.

The aim of this paper is thus to provide an extended set of graphical criteria for the identification of direct causal effects in linear SEMs. We will show examples where the desired causal effect can be identified using algebraic methods but cannot be identified by current graphical methods. Then based on the algebra involved in solving for these causal effects, we will propose this new set of graphical criteria, where we use descendants of either the cause variable or the effect variable as "path-specific instrumental variables", to compute the desired causal effect using observed covariances. This new set of graphical criteria are based on edge removal and the existing graphical criteria of instrumental variables, and the identifiability of certain other total effects, and thus can be easily checked.

The outline of this paper is as follows. First, we provide the preliminary definitions for this paper, including linear SEMs, graphical models, statistical terms such as covariances, Wright's method of analysis, and the identification problem. Then, we recap previous results in solving the parameter identification problem, most notably instrumental variables. Next, we present our new results for the identification of direct causal effects in linear SEMs, which are able to identify direct causal effects which are not possible using previous results based on graphical criteria. Finally, we conclude our paper with a few discussion topics.

2 Preliminaries

In statistical causal analysis, a *directed acyclic graph* (DAG) that represents cause-effect relationships is called a *path diagram*. A *directed graph* is a pair $G = (\boldsymbol{V}, \boldsymbol{E})$, where $\boldsymbol{V}$ is a finite set of vertices and the set $\boldsymbol{E}$ of directed edges is a subset of the set $\boldsymbol{V} \times \boldsymbol{V}$ of ordered pairs of distinct vertices. Regarding the graph theoretic terminology used in this paper, for example, refer to Pearl (Pearl 2009) and Spirtes et al. (Spirtes et al. 2000).

Suppose a DAG $G = (\boldsymbol{V}, \boldsymbol{E})$ with the set $\boldsymbol{V} = \{V_1, \cdots, V_n\}$ of nodes is given. The graph G is called a path diagram, when each child-parent family in G represents a linear structural equation model (SEM):

$$V_i = \sum_{V_j \in \mathrm{pa}(V_i)} \alpha_{v_i v_j} V_j + \epsilon_{v_i}, \qquad i = 1, \ldots, n, \quad (1)$$

where $\mathrm{pa}(V_i)$ is a set of parents of V_i and $\alpha_{v_i v_j} (\neq 0)$ is called a *direct causal effect*. In addition random disturbances $\epsilon_{v_1}, \ldots, \epsilon_{v_n}$ are assumed to be normally distributed with mean 0. Here, when ϵ_{v_i} is correlated with ϵ_{v_j} $(i \neq j)$, this relationship is represented by a bi-directed (dashed) arc between X_i and X_j in G.

Given a path diagram G, we define a *path* between the nodes X and Y as a sequence of vertices, $(V_0 = X, \ldots, V_n = Y)$, where there is an edge between each V_i and V_{i+1} and each vertex appears only once in the sequence. A path is a *directed path* from X to Y if all edges between V_i and V_{i+1} are directed edges from V_i to V_{i+1}. If there is a directed path from X to Y, we say X is an ancestor of Y, $X \in Anc(Y)$, and Y is a descendent of X, $Y \in Desc(X)$. In a directed acyclic graph, we have $Anc(X) \cap Desc(X) = \emptyset$ for every X.

We say that V_i is a *collider* in the path if both the edges between V_{i-1} and V_i and between V_i and V_{i+1} point into V_i. If there are no colliders in the path, we say that the path is an *unblocked path* between X and Y. Given a set of vertices $\boldsymbol{Z}$, we say the path is an *open path* if for all vertices V_i which are not colliders on the path, $V_i \notin \boldsymbol{Z}$, and for all vertices V_i which are colliders on the path, $V \in \boldsymbol{Z}$ where $V = V_i$ or V is a descendant of V_i. If there are no open paths between X and Y given $\boldsymbol{Z}$, we say that X and Y are *d-separated* given $\boldsymbol{Z}$. Otherwise, we say that they are *d-connected*.

The conditional independence induced from a set of equations in the form of Equation 1 can be obtained from the path diagram G according to d-separation (Pearl 2009), that is, when $\boldsymbol{Z}$ d-separates X from Y in G, X is conditionally independent of Y given $\boldsymbol{Z}$ in the corresponding linear SEM (Spirtes et al. 2000). In this paper, it is assumed that a path diagram G and the corresponding linear SEM are faithful to each other; that is, the conditional independence relationships in the linear SEM are also reflected in G, and vice versa (Spirtes et al. 2000).

For further discussion, we denote some notations. Let $\sigma_{xy \cdot z} = \mathrm{cov}(X, Y | \boldsymbol{Z} = \boldsymbol{z})$, $\sigma_{yy \cdot z} = \mathrm{var}(Y | \boldsymbol{Z} = \boldsymbol{z})$ and $\beta_{yx \cdot z} = \sigma_{xy \cdot z} / \sigma_{xx \cdot z}$ be a *conditional covariance* between X and Y given $\boldsymbol{Z} = \boldsymbol{z}$, a *conditional variance* of Y given $\boldsymbol{Z} = \boldsymbol{z}$ and the *regression coefficient* of x in the regression model of Y on x and $\boldsymbol{z}$, respectively. When $\boldsymbol{Z}$ is an empty set, they are omitted from these arguments.

A *total effect* τ_{yx} of X on Y is defined as the total sum of the products of the direct causal effects on the sequence of directed edges along all directed paths from X to Y. In addition, $\gamma_{yx} = \sigma_{xy} - \tau_{yx}$ is called a *spurious correlation* between X and Y.

A *path-specific total effect* is defined as the total sum of the product of the direct causal effects on the sequence of directed edges along directed paths of our interests from X to Y. For example, we define $\tau_{yx \cdot z}$ as the path-specific effects where all paths that pass through any variable in $\boldsymbol{Z}$ are not counted. Similar terms such as a *path-specific correlation* and a *path-specific spurious correlation* can be defined similarly.

Wright's method of path analysis (Wright 1934), which plays an important role in this paper, can be used to compute the covariance of two variables X and Y given a path diagram G. If the set S contains all paths $path = (V_0 = X, V_1, \ldots, V_n = Y)$ that are unblocked paths between X and Y, we have:

$$\sigma_{xy} = \sum_{path} p_{v_0 v_0} \prod_{i=0,\ldots,n-1} p_{v_i v_{i+1}}, \quad (2)$$

where $p_{v_i v_{i+1}}$ is the parameter of the edge between V_i

and V_{i+1}, which is either $\alpha_{v_{i+1}v_i}$ (or $\alpha_{v_i v_{i+1}}$) if it is a directed edge, or $\gamma_{v_{i+1}v_i}$ if it is a bi-directed edge. We define $p_{v_0v_0}$ as $\sigma_{v_0v_0}$ if all edges in *path* are directed edges, or 1 otherwise.

Given the matrix of observed covariances Σ, we say that a causal parameter, such as a total effect and a direct causal effect, is *identified* if there is a unique solution of this parameter given the covariances. If all direct causal effects can be identified, we say that the model is identified.

The *single-door criterion* is one of the famous graphical identification conditions for the direct causal effects, that is, the direct causal effect α_{yx} of X on Y is identifiable and is equal to $\beta_{yx\cdot z}$, if there exist a set $\boldsymbol{Z}$ of variables such that $\boldsymbol{Z}$ contains no descendant of Y, and $\boldsymbol{Z}$ d-separates X from Y in $G_{X \to Y}$, formed by removing $X \to Y$ from the path diagram G (Pearl 2009). A set $\boldsymbol{Z}$ of variable satisfying both (i) and (ii) is said to satisfy the single-door criterion relative to (X, Y).

3 Previous Results on Parameter Identification

There have been many work done on the problems of model identification and parameter identification using graphical test (Pearl 2009; Brito and Pearl 2002a,b,c; Tian 2004, 2005). Here we will focus only on the problem of parameter identification, in particular the identification of direct causal effects.

The previous most general result for the graphical identification of direct causal effects is the use of an *IV-pair* (Brito and Pearl 2002c), which embraces both *instrumental variables* (Bowden and Turkington 1984) and regression methods.

Lemma 1 *Given a path diagram G which contains the directed edge $X \to Y$, we say that a variable W is an instrumental variable for $X \to Y$ given $\boldsymbol{Z}$, a (possibly empty) set of variables which does not contain any variable from W, X, Y, or $Desc(Y)$, if the following two conditions are satisfied:*

1. *In the path diagram G, W and X are d-connected given $\boldsymbol{Z}$, or $W = X$.*
2. *In the path diagram $G_{\backslash X \to Y}$, formed by removing $X \to Y$ from G, W and Y are d-separated given $\boldsymbol{Z}$;*

Then, the direct causal effect α_{yx} is given by:

$$\alpha_{yx} = \frac{\sigma_{wy\cdot z}}{\sigma_{wx\cdot z}}. \quad (3)$$

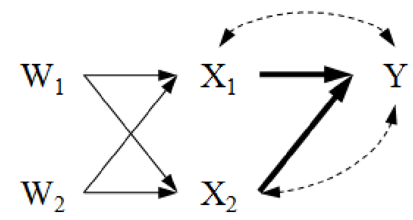

Figure 1: A path diagram where a multiple IV-pair are necessary to identify the direct causal effects of X_1 (or X_2) to Y.

The pair $(W, \boldsymbol{Z})$ can also be called an IV-pair for $X \to Y$. Some identifiable direct causal effects cannot be found using a single IV-pair, but by the collective action of a multiple IV-pair, such as the example in Figure 1. We now define the conditions where a set of direct causal effects can be identified using multiple IV-pairs. The following lemma is adapted from previous work, where the multiple IV-pair are called *accessory sets* (Tian 2007b).

Lemma 2 *Given a path diagram G which contains the directed edges $X_1 \to Y, \ldots, X_k \to Y$, we say that a set of variables $W_1, \ldots, W_k$ are instrumental variables for $X_1 \to Y, \ldots, X_k \to Y$ given $\boldsymbol{Z}$, a (possibly empty) set of variables which does not contain any variable from $W_1, \ldots, W_k$, $X_1, \ldots, X_k$, Y, or $Desc(Y)$, if the following two conditions are satisfied:*

1. *In the path diagram G, each pair of W_i and X_i are d-connected given $\boldsymbol{Z}$, or $W_i = X_i$. Moreover, for any $i \neq j$, there must be exist an open path between W_i and X_i given $\boldsymbol{Z}$ (denoted $path_i$), and an open path between W_j and X_j given $\boldsymbol{Z}$ (denoted $path_j$), such that either they do not share any common variable, or if they share a common variable V, either one of the following is true (but not both):*
   - *Both the sub-path of $path_i$ between W_i and U, and the sub-path of $path_j$ between U and X_j, point into U;*
   - *Both the sub-path of $path_i$ between U and W_j, and the sub-path of $path_i$ between W_j and U, point into U.*

   This criterion is called the G criterion *(Brito and Pearl 2006), and guarantees that the system of equations we use to solve for the parameters $\alpha_{yx_1}, \ldots, \alpha_{yx_k}$ are linearly independent.*
2. *In the path diagram $G_{\backslash X_1 \to Y, \ldots, X_k \to Y}$, formed by removing $X_1 \to Y, \ldots, X_k \to Y$ from G, $\{W_1, \ldots, W_k\}$ and Y are d-separated given $\boldsymbol{Z}$;*

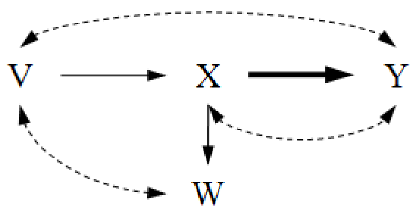

Figure 2: Path diagram for Example 1.

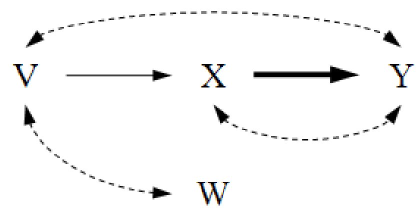

Figure 3: Removing all directed paths from X to W from the path diagram of Figure 2. W now satisfies the instrumental variable condition.

Then, the direct causal effects $\alpha_{yx_1}, \ldots, \alpha_{yx_k}$ can be solved by a system of k equations:

$$\sigma_{w_i y \cdot z} = \sum_{j=1,\ldots,k} \sigma_{w_i x_j \cdot z} \alpha_{yx_j}, \qquad i = 1, \ldots, k.$$

However, there are many cases where a single direct causal effect is identifiable even though neither IV-pair nor a set of variables satisfying the single-door criterion can be found. We will illustrate this in the next section with a few examples, and extend previous results on graphical identification to find the direct causal effects.

4 New Results on Parameter Identification

Our results are based on the following. Based on whether W is a descendant of X or Y, we will remove certain edges from the path diagram G to obtain G', where W satisfies the instrumental variable condition in G'. This means the direct causal effect α_{yx} can be computed using the matrix of covariance values Σ' of G'. We then relate Σ' with Σ, the matrix of covariance values of G, under certain restrictions, in order to compute the direct causal effect α_{yx} from Σ.

4.1 Descendants of Cause Variable

Example 1 *Given the path diagram shown in Figure 2 and its corresponding linear SEM, the direct causal effects of α_{xv} and α_{wx} can be identified by the single-door criterion (Pearl 2009):*

$$\alpha_{xv} = \beta_{xv}, \qquad \alpha_{wx} = \beta_{wx\cdot v}.$$

We now want to find the direct causal effect of α_{yx}. However, no IV-pair can be used to find α_{yx} using Lemma 1. Instead, we use Wright's method of path analysis (Equation 2). In particular, we have:

$$\begin{aligned} \sigma_{wy} &= \alpha_{xv}\alpha_{yx}\gamma_{wv} + \alpha_{wx}(\alpha_{xv}\gamma_{yv} + \alpha_{yx}\sigma_{xx} + \gamma_{yx}), \\ \sigma_{wx} &= \alpha_{xv}\gamma_{wv} + \alpha_{wx}\sigma_{xx}, \\ \sigma_{yx} &= \alpha_{xv}\gamma_{yv} + \alpha_{yx}\sigma_{xx} + \gamma_{yx}. \end{aligned}$$

Therefore, $\sigma_{wy} - \beta_{wx\cdot v}\sigma_{yx} = \alpha_{xv}\alpha_{yx}\gamma_{wv}$, and $\sigma_{wx} - \beta_{wx\cdot v}\sigma_{xx} = \alpha_{xv}\gamma_{wv}$, and α_{yx} can be computed by:

$$\alpha_{yx} = \frac{\sigma_{wy} - \beta_{wx\cdot v}\sigma_{yx}}{\sigma_{wx} - \beta_{wx\cdot v}\sigma_{xx}}.$$

Notice that in this example, neither W nor V can be used as an instrumental variable to identify α_{yx}. However, if we consider a latent variable U along the bi-directed edge between W and V, this latent variable, if observable, can be used as an instrumental variable to identify α_{wx} (Cai and Kuroki 2008), meaning that α_{yx} can be identified if we can "indirectly" estimate the correlations both between U and X, and U and Y. The variable W, as an "descendant" of U, can potentially be used so, except that it is also a descendant of X, and the presence of an open path $W \leftarrow X \leftrightarrow Y$ in G makes it invalid to be used as an instrumental variable. Therefore, we have to consider the path-specific correlations between W and Y and between W and X, only through the bi-directed edge between W and V, while discounting those through the total effects from X to W. We first make the following definition.

Definition 1 *Given a path diagram G which contains the directed edge $X \rightarrow Y$, and a variable $W \in Desc(X), \notin Desc(Y)$, we define the path diagram $G' = G_{\setminus X\rightarrow\{W\}\cup Anc(W)}$ as formed by removing all directed edges $X \rightarrow V$, where $V \in \{W\} \cup Anc(W)$.*

The effect of removing the set of directed edges $X \rightarrow \{W\} \cup Anc(W)$ from G is to remove all directed paths from X to W in G, meaning there is no total effect of X on W in $G_{\setminus X\rightarrow\{W\}\cup Anc(W)}$. For example, the path diagram in Figure 2 is transformed into Figure 3 by removing the directed edge $X \rightarrow W$, so as to remove all directed paths from X to W. We notice that in this new path diagram, W can be used as an instrumental variable to identify $X \rightarrow Y$. The question now becomes: what are the relations between the matrices of covariance values in G and $G_{\setminus X\rightarrow\{W\}\cup Anc(W)}$, denoted as Σ and Σ' respectively?

Before we find these relations, we have to make sure that all relevant paths in G for the computation of α_{yx}, i.e., the path-specific correlations between W and Y given $\boldsymbol{Z}$, and the path-specific correlations between W and X given $\boldsymbol{Z}$, are preserved in $G_{\backslash X \to \{W\} \cup Anc(W)}$. This can be done by putting a further restriction on $\boldsymbol{Z}$, besides that it does not contain any variable from W, X, Y, or $Desc(Y)$. For any directed edge $X \to V$ removed from G to form $G_{\backslash X \to \{W\} \cup Anc(W)}$, $\boldsymbol{Z}$ cannot contain V or any descendant of V. Otherwise, an unblocked path from X to $Z \in \boldsymbol{Z}$ in G will not be in $G_{\backslash X \to \{W\} \cup Anc(W)}$ after removing $X \to V$ from G. This restriction on $\boldsymbol{Z}$ is equivalent to restricting that $\boldsymbol{Z}$ does not contain any variable from W, $Desc(W)$, or $Anc(W) \cap Desc(X)$.

Assuming that W is an instrumental variable for $X \to Y$ given $\boldsymbol{Z}$ in $G_{\backslash X \to \{W\} \cup Anc(W)}$, we have the following relations between G and $G_{\backslash X \to \{W\} \cup Anc(W)}$:

- The sets of unblocked paths between X and Y in G and $G_{\backslash X \to \{W\} \cup Anc(W)}$ are the same. This is because for any V where the directed edge $X \to V$ is removed from G, there is no directed path $X \to V \longrightarrow Y$ in G. Otherwise, there would be a directed path $V \longrightarrow Y$ in both G and $G_{\backslash X \to \{W\} \cup Anc(W)}$, meaning that there would be an unblocked path $W \longleftarrow V \longrightarrow Y$ in $G_{\backslash X \to \{W\} \cup Anc(W)}$, violating the instrumental variable condition of W (note that no variable $Z \in \boldsymbol{Z}$ can block this path as all variables in this path are descendants of V).

- For any variable $Z \in \boldsymbol{Z}$, the sets of unblocked paths between X and Z in G and $G_{\backslash X \to \{W\} \cup Anc(W)}$ are the same, as stated above due to the restriction on $\boldsymbol{Z}$. Moreover, because the sets of unblocked paths between X and Y in G and $G_{\backslash X \to \{W\} \cup Anc(W)}$ are the same, the sets of unblocked paths between Y and Z in G and $G_{\backslash X \to \{W\} \cup Anc(W)}$ are also the same. Therefore, we have:

$$\begin{aligned} \sigma_{xy} &= \sigma'_{xy}; \\ \sigma_{xz} &= \sigma'_{xz}; \\ \sigma_{yz} &= \sigma'_{yz}. \end{aligned}$$

- By transforming from G to $G_{\backslash X \to \{W\} \cup Anc(W)}$, we remove all directed paths from X to W. Because the sets of unblocked paths between X and Y and between X and Z in G and G' are the same, we have:

$$\begin{aligned} \sigma_{wx} &= \sigma'_{wx} + \tau_{wx}\sigma_{xx}; \\ \sigma_{wy} &= \sigma'_{wy} + \tau_{wx}\sigma_{xy}; \\ \sigma_{wz} &= \sigma'_{wz} + \tau_{wx}\sigma_{xz}, \end{aligned}$$

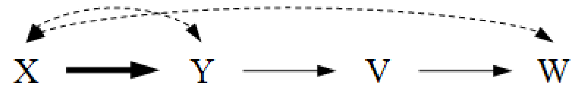

Figure 4: Path diagram for Example 2.

where τ_{wx} is the total effect of X on W in G.

After establishing these relations, we are ready to compute the causal effects α_{yx}. By Lemma 1, we have:

$$\begin{aligned} &\alpha_{yx} \\ &= \frac{\sigma'_{wy \cdot z}}{\sigma'_{wx \cdot z}} \\ &= \frac{\sigma'_{wy} - \sigma'_{wz}\sigma'_{yz}/\sigma'_{zz}}{\sigma'_{wx} - \sigma'_{wz}\sigma'_{xz}/\sigma'_{zz}} \\ &= \frac{(\sigma_{wy} - \tau_{wx}\sigma_{xy}) - (\sigma_{wz} - \tau_{wx}\sigma_{xz})\sigma_{yz}/\sigma_{zz}}{(\sigma_{wx} - \tau_{wx}\sigma_{xx}) - (\sigma_{wz} - \tau_{wx}\sigma_{xz})\sigma_{xz}/\sigma_{zz}} \\ &= \frac{(\sigma_{wy} - \sigma_{wz}\sigma_{yz}/\sigma_{zz}) - \tau_{wx}(\sigma_{xy} - \sigma_{xz}\sigma_{yz}/\sigma_{zz})}{(\sigma_{wx} - \sigma_{wz}\sigma_{xz}/\sigma_{zz}) - \tau_{wx}(\sigma_{xx} - \sigma_{xz}\sigma_{xz}/\sigma_{zz})} \\ &= \frac{\sigma_{wy \cdot z} - \tau_{wx}\sigma_{xy \cdot z}}{\sigma_{wx \cdot z} - \tau_{wx}\sigma_{xx \cdot z}}. \end{aligned}$$

Therefore, we have the following theorem.

Theorem 1 *If $W \in Desc(X), \notin Desc(Y)$ is an instrumental variable for $X \to Y$ given $\boldsymbol{Z}$, a (possibly empty) set of variables which does not contain any variable from W, X, Y, $Desc(Y)$, $Desc(W)$, or $Anc(W) \cap Desc(X)$, in the path diagram $G_{\backslash X \to \{W\} \cup Anc(W)}$ (Definition 1), the direct causal effect α_{yx} is given by:*

$$\alpha_{yx} = \frac{\sigma_{wy \cdot z} - \tau_{wx}\sigma_{xy \cdot z}}{\sigma_{wx \cdot z} - \tau_{wx}\sigma_{xx \cdot z}}. \quad (4)$$

Equation 4 expands on the previous use of the IV-pair (Equation 3), by allowing the use of $W \in Desc(X), \notin Desc(Y)$ as a "path-specific instrumental variable", showing that the direct causal effect α_{yx} is identifiable as long as the total effect of X on W, τ_{wx}, is identifiable. For example, as the path diagram in Figure 2 satisfies the conditions of Theorem 1, we can use Equation 4 to compute α_{yx}. In particular, the total effect of X on W is given by $\tau_{wx} = \alpha_{wx} = \beta_{wx \cdot v}$, and this yields the result in Example 1.

4.2 Descendants of Effect Variable

Example 2 *Given the path diagram shown in Figure 4 and its corresponding linear SEM,*[1] *the direct*

[1] This model is shown to be identifiable (Brito and Pearl 2006), and is called a *P-structure* (Tian 2007a).

causal effects α_{vy} and α_{wv} can be identified by the single-door criterion (Pearl 2009):

$$\alpha_{vy} = \beta_{vy}, \quad \alpha_{wv} = \beta_{wv\cdot y}.$$

We now want to find the direct causal effect of α_{yx}. However, no IV-pair can be used to find α_{yx} using Lemma 1. Instead, we use Wright's method of path analysis (Equation 2). In particular, we have:

$$\begin{aligned}\sigma_{yx} &= \alpha_{yx}\sigma_{xx} + \gamma_{yx},\\ \sigma_{wx} &= \alpha_{wv}\alpha_{vy}(\alpha_{yx}\sigma_{xx} + \gamma_{yx}) + \gamma_{wx},\\ \sigma_{wy} &= \alpha_{wv}\alpha_{vy}\sigma_{yy} + \alpha_{yx}\gamma_{wx}.\end{aligned}$$

Therefore, $\sigma_{wy} - \beta_{wv\cdot y}\beta_{vy}\sigma_{yy} = \alpha_{yx}\gamma_{wx}$, and $\sigma_{wx} - \beta_{wv\cdot y}\beta_{vy}\sigma_{yx} = \gamma_{wx}$, and α_{yx} can be computed by:

$$\alpha_{yx} = \frac{\sigma_{wy} - \beta_{wv\cdot y}\beta_{vy}\sigma_{yy}}{\sigma_{wx} - \beta_{wv\cdot y}\beta_{vy}\sigma_{yx}}.$$

Notice that in this example, W cannot be used as an instrumental variable to identify α_{yx} (and we cannot use V to block the path between W and Y since V is a descendant of Y). However, if we consider a latent variable U along the bi-directed edge between W and X, this latent variable, if observable, can be used as an instrumental variable to identify α_{wx} (Cai and Kuroki 2008), meaning that α_{yx} can be identified if we can "indirectly" estimate the correlations both between U and X, and U and Y. The variable W, as an "descendant" of U, can potentially be used so, except that it is also a descendant of Y, and the presence of an open path $Y \to V \leftrightarrow W$ in G makes it invalid to be used as an instrumental variable. Therefore, we have to consider the path-specific correlations between W and Y and between W and X, only through the bi-directed edge between W and X, while discounting those through the total effects from Y to W. We first make the following definition.

Definition 2 *Given a path diagram G which contains the directed edge $X \to Y$, and a variable $W \in Desc(Y)$, we define the path diagram $G' = G_{\backslash Y\to\{W\}\cup Anc(W)}$ as formed by removing all directed edges $Y \to V$, where $V \in \{W\} \cup Anc(W)$.*

The effect of removing the set of directed edges $Y \to \{W\} \cup Anc(W)$ from G is to remove all directed paths from Y to W in G, meaning there is no total effect of Y on W in $G_{\backslash Y\to\{W\}\cup Anc(W)}$. For example, the path diagram in Figure 4 is transformed into Figure 5 by removing the directed edge $Y \to V$, so as to remove all directed paths from Y to W. We notice that in this new path diagram, W can be used as an instrumental variable to identify $X \to Y$. The question now becomes: what are the relations between the matrices of covariance values in G and $G_{\backslash Y\to\{W\}\cup Anc(W)}$, denoted as Σ and Σ' respectively?

Figure 5: Removing all directed paths from Y to W from the path diagram of Figure 4. W now satisfies the instrumental variable condition.

Before we find these relations, we have to make sure that all relevant paths in G for the computation of α_{yx}, i.e., the path-specific correlations between W and Y given $\boldsymbol{Z}$, and the path-specific correlations between W and X given $\boldsymbol{Z}$, are preserved in $G_{\backslash Y\to\{W\}\cup Anc(W)}$. The original restriction on $\boldsymbol{Z}$, that it does not contain any variable from W, X, Y, or $Desc(Y)$, is sufficient, since for any directed edge $X \to V$ removed from G to form $G_{\backslash Y\to\{W\}\cup Anc(W)}$, $\boldsymbol{Z}$ cannot contain V or any descendant of V.

Assuming that W is an instrumental variable for $X \to Y$ given $\boldsymbol{Z}$ in $G_{\backslash Y\to\{W\}\cup Anc(W)}$, we have the following relations between G and $G_{\backslash Y\to\{W\}\cup Anc(W)}$:

- The sets of unblocked paths between X and Y in G and $G_{\backslash Y\to\{W\}\cup Anc(W)}$ are the same, and for any variable $Z \in \boldsymbol{Z}$, the sets of unblocked paths between X and Z and between Y and Z in G and $G_{\backslash Y\to\{W\}\cup Anc(W)}$ are the same. This is because for any V where the directed edge $X \to V$ is removed from G, $V \neq X, Y, Z$, nor can V be an ancestor of X, Y, or Z. Therefore, we have:

$$\begin{aligned}\sigma_{xy} &= \sigma'_{xy};\\ \sigma_{xz} &= \sigma'_{xz};\\ \sigma_{yz} &= \sigma'_{yz}.\end{aligned}$$

- By transforming from G to $G_{\backslash Y\to\{W\}\cup Anc(W)}$, we remove all directed paths from Y to W. Because the sets of unblocked paths between X and Y and between Y and Z in G and $G_{\backslash Y\to\{W\}\cup Anc(W)}$ are the same, we have:

$$\begin{aligned}\sigma_{wx} &= \sigma'_{wx} + \tau_{wy}\sigma_{xy};\\ \sigma_{wy} &= \sigma'_{wy} + \tau_{wy}\sigma_{yy};\\ \sigma_{wz} &= \sigma'_{wz} + \tau_{wy}\sigma_{yz}.\end{aligned}$$

After establishing these relations, we are ready to compute the causal effects α_{yx}. By Lemma 1, we have:

$$\alpha_{yx} = \frac{\sigma'_{wy\cdot z}}{\sigma'_{wx\cdot z}}$$

$$= \frac{\sigma'_{wy} - \sigma'_{wz}\sigma'_{yz}/\sigma'_{zz}}{\sigma'_{wx} - \sigma'_{wz}\sigma'_{xz}/\sigma'_{zz}}$$
$$= \frac{(\sigma_{wy} - \tau_{wy}\sigma_{yy}) - (\sigma_{wz} - \tau_{wy}\sigma_{yz})\sigma_{yz}/\sigma_{zz}}{(\sigma_{wx} - \tau_{wy}\sigma_{xy}) - (\sigma_{wz} - \tau_{wy}\sigma_{yz})\sigma_{xz}/\sigma_{zz}}$$
$$= \frac{(\sigma_{wy} - \sigma_{wz}\sigma_{yz}/\sigma_{zz}) - \tau_{wy}(\sigma_{yy} - \sigma_{yz}\sigma_{yz}/\sigma_{zz})}{(\sigma_{wx} - \sigma_{wz}\sigma_{xz}/\sigma_{zz}) - \tau_{wx}(\sigma_{xy} - \sigma_{xz}\sigma_{yz}/\sigma_{zz})}$$
$$= \frac{\sigma_{wy\cdot z} - \tau_{wy}\sigma_{yy\cdot z}}{\sigma_{wx\cdot z} - \tau_{wy}\sigma_{xy\cdot z}}.$$

Therefore, we have the following theorem.

Theorem 2 *If $W \in Desc(Y)$ is an instrumental variable for $X \to Y$ given $\boldsymbol{Z}$, a (possibly empty) set of variables which does not contain any variable from W, X, Y, or $Desc(Y)$, in the path diagram $G_{\backslash Y \to \{W\} \cup Anc(W)}$ (Definition 2), the direct causal effect α_{yx} is given by:*

$$\alpha_{yx} = \frac{\sigma_{wy\cdot z} - \tau_{wy}\sigma_{yy\cdot z}}{\sigma_{wx\cdot z} - \tau_{wy}\sigma_{xy\cdot z}}. \quad (5)$$

Equation 5 expands on the previous use of the IV-pair (Equation 3), by allowing the use of $W \in Desc(Y)$ as a "path-specific instrumental variable", showing that the direct causal effect α_{yx} is identifiable as long as the total effect of Y on W, τ_{wy}, is identifiable. For example, as the path diagram in Figure 4 satisfies the conditions of Theorem 2, we can use Equation 5 to compute α_{yx}. In particular, the total effect of X on W is given by $\tau_{wy} = \beta_{wv\cdot y}\beta_{vy}$, and this yields the result in Example 2.

It is possible that given $W \in Desc(Y)$, two stages of edge removal can be applied such that W becomes an instrumental variable for $X \to Y$ given $\boldsymbol{Z}$ in the resulting path diagram. First, we form the path diagram $G' = G_{\backslash Y \to \{W\} \cup Anc(W)}$ by removing all directed edges $Y \to V$, where $V = W$ or $V \in Anc(W)$. Then, we form the path diagram $G'' = G_{\backslash Y \to \{W\} \cup Anc(W), X \to \{W\} \cup Anc'(W)}$ by removing all directed edges $X \to V'$, where $V' = W$ or $V' \in Anc'(W)$, where $Anc'(W)$ are the ancestors of W in $G' = G_{\backslash Y \to \{W\} \cup Anc(W)}$. The restriction on $\boldsymbol{Z}$ is that does not contain any variable from W, X, Y, $Desc(Y)$, $Desc(W)$, or $Anc'(W) \cap Desc(X)$. By Theorem 1, we have:

$$\alpha_{yx} = \frac{\sigma'_{wy\cdot z} - \tau'_{wx}\sigma'_{xy\cdot z}}{\sigma'_{wx\cdot z} - \tau'_{wx}\sigma'_{xx\cdot z}},$$

where Σ' is the matrix of covariance values in $G_{\backslash Y \to \{W\} \cup Anc(W)}$, and τ'_{wx} is the total effect of X on W in $G_{\backslash Y \to \{W\} \cup Anc(W)}$. Since we remove all directed paths from Y to W to obtain $G_{\backslash Y \to \{W\} \cup Anc(W)}$ from G, we have $\tau'_{wx} = \tau_{wx\cdot y}$. Moreover, we have $\sigma'_{xy\cdot z} = \sigma_{xy\cdot z}$ and $\sigma'_{xx\cdot z} = \sigma_{xx\cdot z}$. Therefore, we have:

$$\alpha_{yx} = \frac{\sigma_{wy\cdot z} - \tau_{wy}\sigma_{yy\cdot z} - \tau_{wx\cdot y}\sigma_{xy\cdot z}}{\sigma_{wx\cdot z} - \tau_{wy}\sigma_{xy\cdot z} - \tau_{wx\cdot y}\sigma_{xx\cdot z}}.$$

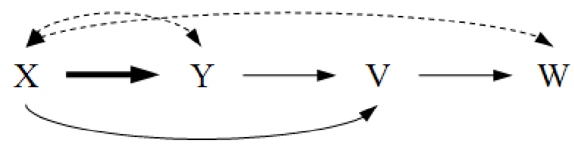

Figure 6: Path diagram which satisfies the conditions of Theorem 3.

Therefore, we have the following theorem.

Theorem 3 *If $W \in Desc(Y)$ is an instrumental variable for $X \to Y$ given $\boldsymbol{Z}$, a (possibly empty) set of variables which does not contain any variable from W, X, Y, $Desc(Y)$, $Desc(W)$, or $Anc'(W) \cap Desc(X)$ ($Anc'(W)$ are the ancestors of W in $G_{\backslash Y \to \{W\} \cup Anc(W)}$), in the path diagram $G_{\backslash Y \to \{W\} \cup Anc(W), X \to \{W\} \cup Anc'(W)}$ (Definitions 1 and 2), the direct causal effect α_{yx} is given by:*

$$\alpha_{yx} = \frac{\sigma_{wy\cdot z} - \tau_{wy}\sigma_{yy\cdot z} - \tau_{wx\cdot y}\sigma_{xy\cdot z}}{\sigma_{wx\cdot z} - \tau_{wy}\sigma_{xy\cdot z} - \tau_{wx\cdot y}\sigma_{xx\cdot z}}. \quad (6)$$

For an example, in the path diagram in Figure 6 (adapted from the path diagram in Figure 4 by adding a directed edge $X \to V$), the total effect of Y on W is given by $\tau_{wy} = \beta_{wv\cdot yx}\beta_{vy\cdot x}$, and the total effect of X on W given Y is given by $\tau_{wx\cdot y} = \beta_{wv\cdot yx}\beta_{vx\cdot y}$, and we can use Equation 6 to compute α_{yx}.

Finally, the results in Theorems 1, 2 and 3 can all be extended to identify multiple direct causal effects similar to Lemma 2, as long as the G criterion is satisfied. However, we do not state the details here.

5 Discussion and Conclusion

In this paper, we presented new results for the identification of direct causal effects in linear SEMs based on graphical criteria after edge removal, which are able to identify direct causal effects which are not possible using previous results based on graphical criteria. The results allow us to identify the direct causal effect, by first checking whether the model satisfies the graphical criteria, then using the necessary covariance values and other total effects to compute the direct causal effect, instead of solving the whole set of equations for all variables given by Wright's method of analysis. As our new results are based on edge removal and the determination of whether W can be used as an instrumental variable in the resulting path diagram, existing algorithms can be used to check for the satisfying conditions of our theorems. While the identification of certain other total effects are also necessary, they can also be identified using existing graphical methods.

Even in models where previous graphical criteria can be used to identify a certain causal effect, our new set of criteria may provide another distinct function for computing this causal effect. The concept of *robustness* (Pearl 2004) deals with whether a function for computing a causal effect is still valid when certain independence relations are relaxed in a model (by adding edges between variables). If for all super-graphs of our model, function A is valid whenever function B is valid, we say that function A is at least as robust as function B, and should at least be more preferred than function B, because function A will remain valid even in cases where function B is no longer valid when some of our current independence relations are relaxed. Moreover, if we are given two functions that are no more robust than one another, then the computation of the direct causal effect using these two different functions (and the fact that the two computations agree) will greatly confirm the correctness of our model.

It remains to be seen if the graphical criteria given in this paper, combined with previous methods such as IV-pairs and accessory sets, are complete, i.e., they are necessary conditions for the identification of causal effects in linear SEMs.

Acknowledgments

The research work of Hei Chan was performed while employed at the National Institute of Advanced Industrial Science and Technology (AIST).

The research work of Manabu Kuroki was supported by the Ministry of Education, Culture, Sports, Science and Technology of Japan, the Mazda Foundation and the JIST Foundation.

References

Kenneth A. Bollen. *Structural Equations with Latent Variables.* John Wiley, 1989.

Roger J. Bowden and Darrell A. Turkington. *Instrumental Variables.* Cambridge University Press, 1984.

Carlos Brito and Judea Pearl. A graphical criterion for the identification of causal effects in linear models. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI)*, pages 533–538. AAAI Press, 2002a.

Carlos Brito and Judea Pearl. A new identification condition for recursive models with correlated errors. *Structural Equation Modeling*, 9(4):459–474, 2002b.

Carlos Brito and Judea Pearl. Generalized instrumental variables. In *Proceedings of the Eighteenth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 85–93. Morgan Kaufmann Publishers, 2002c.

Carlos Brito and Judea Pearl. Graphical condition for identification in recursive SEM. In *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 47–54. AUAI Press, 2006.

Zhihong Cai and Manabu Kuroki. On identifying total effects in the presence of latent variables and selection bias. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 62–69. AUAI Press, 2008.

Otis D. Duncan. *Introduction to Structural Equation Models.* Academic Press, 1975.

Franklin M. Fisher. *The Identification Problem in Econometrics.* McGraw-Hill, 1966.

Judea Pearl. Robustness of causal claims. In *Proceedings of the Twentieth Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 446–453. AUAI Press, 2004.

Judea Pearl. *Causality: Models, Reasoning, and Inference.* Cambridge University Press, 2nd edition, 2009. 1st edition, 2000.

Peter Spirtes, Clark Glymour, and Richard Scheines. *Causation, Prediction, and Search.* MIT Press, 2000.

Jin Tian. Identifying linear causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence (AAAI)*, pages 104–111. AAAI Press, 2004.

Jin Tian. Identifying direct causal effects in linear models. In *Proceedings of the Nineteenth National Conference on Artificial Intelligence (AAAI)*, pages 346–353. AAAI Press, 2005.

Jin Tian. On the identification of a class of linear models. In *Proceedings of the Twenty-Second National Conference on Artificial Intelligence (AAAI)*, pages 1284–1289. AAAI Press, 2007a.

Jin Tian. A criterion for parameter identification in structural equation models. In *Proceedings of the Twenty-Third Conference on Uncertainty in Artificial Intelligence (UAI)*, pages 392–399. AUAI Press, 2007b.

Sewall Wright. The method of path coefficients. *Annals of Mathematical Science*, 5(3):161–215, 1934.