

Judea Pearl\* and Azaria Paz

# Confounding Equivalence in Causal Inference

**Abstract:** The paper provides a simple test for deciding, from a given causal diagram, whether two sets of variables have the same bias-reducing potential under adjustment. The test requires that one of the following two conditions holds: either (1) both sets are admissible (i.e. satisfy the back-door criterion) or (2) the Markov boundaries surrounding the treatment variable are identical in both sets. We further extend the test to include treatment-dependent covariates by broadening the back-door criterion and establishing equivalence of adjustment under selection bias conditions. Applications to covariate selection and model testing are discussed.

**Keywords:** confounding, model testing, selection bias, extended back-door, covariate selection

DOI 10.1515/jci-2013-0020

## 1 Introduction

The common method of estimating causal effects in observational studies is to *adjust* for a set of variables (or “covariates”) judged to be “confounders,” that is, variables capable of producing spurious associations between treatment and outcome, not attributable to their causative dependence. While adjustment tends to reduce the bias produced by such spurious associations, the bias-reducing potential of any set of covariates depends crucially on the causal relationships among all variables affecting treatment or outcome, hidden as well as visible. Such relationships can effectively be represented in the form of directed acyclic graphs (DAGs) [1–5].

Most studies of covariate selection have aimed to define and identify “admissible” sets of covariates, also called “sufficient sets,” namely, a set of covariates that, if adjusted for, would yield asymptotically unbiased estimates of the causal effect of interest [6–8]. A graphical criterion for selecting an admissible set is given by the “back-door” test [8, 9] which was shown to entail zero bias, or “no confoundedness,” assuming correctness of the causal assumptions encoded in the DAG. Related notions are “exchangeability” [6], “exogeneity” [10], and “strong ignorability” [11].

This paper addresses a different question: Given two sets of variables in a DAG, decide if the two are equally valuable for adjustment, namely, whether adjustment for one set is guaranteed to yield the same asymptotic bias as adjustment for the other.

The reasons for posing this question are several. First, an investigator may wish to assess, prior to taking any measurement, whether two candidate sets of covariates, differing substantially in dimensionality, measurement error, cost, or sample variability are equally valuable in their bias-reduction potential. Whenever such equality holds, we say that the two sets are *confounding equivalent* or *c-equivalent* and the statistical condition implied by such equality is called *c-equivalence test*. Second, an investigator may face a post-measurement choice among several statistical estimates each based on a different set of covariates. If the sets are known to be *c-equivalent*, the choice among their estimates can be made on the basis of variance minimization, rather than bias-reduction considerations [12, 13]. Third, assuming that the structure of the underlying DAG is only partially known, one may wish to assess, using *c-equivalence tests*, whether a

---

\*Corresponding author: **Judea Pearl**, Department of Computer Science, University of California – Los Angeles, Los Angeles, CA 90095-1596 USA, E-mail: judea@cs.ucla.edu

**Azaria Paz**, Department of Computer Science, Technion IIT, Haifa 3200, Israel, E-mail: paz@cs.technion.ac.il

given structure is compatible with the data at hand; structures that predict equality of post-adjustment associations must be rejected if, after adjustment, such equality is not found in the data.

In Section 2, we define  $c$ -equivalence and review the auxiliary notions of admissibility,  $d$ -separation, and the back-door criterion. Section 3 derives statistical and graphical conditions for  $c$ -equivalence, the former being sufficient while the latter necessary and sufficient. Section 4 presents a simple algorithm for testing  $c$ -equivalence, assuming that the two sets contain no treatment-dependent variables. Section 5 generalizes this algorithm to any two sets of covariates by extending the back-door criterion to allow treatment-dependent variables. Section 6 gives a statistical interpretation to the graphical test of Section 4, not invoking the causal notion of “admissibility” or “no confoundedness.” Finally, Section 7 demonstrates potential applications of  $c$ -equivalence in effect estimation and model testing.

## 2 Preliminaries: $c$ -equivalence and admissibility

Let  $X, Y$ , and  $Z$  be three disjoint subsets of discrete variables, and  $P(x, y, z)$  their joint distribution. We are concerned with expressions of the type<sup>1</sup>

$$A_P(x, y, Z) = \sum_z P(y|x, z)P(z) \quad (1)$$

Such expressions, which we name “adjustment estimands,” are often used to approximate the causal effect of  $X$  on  $Y$ , where the set  $Z$  is chosen to include variables judged to be “confounders.” By *adjusting* for these variables, one hopes to create conditions that eliminate spurious dependence and thus obtain an unbiased estimate of the causal effect of  $X$  and  $Y$ , written  $P(y|do(x))$  (see Pearl [8, 9] for formal definition and methods of estimation).

### Definition 1. ( $c$ -equivalence)

Define two sets,  $T$  and  $Z$  as  $c$ -equivalent (relative to  $X$  and  $Y$ ), written  $T \sim Z$ , if the following equality holds for every  $x$  and  $y$ :

$$\sum_t P(y|x, t)P(t) = \sum_z P(y|x, z)P(z) \quad \forall x, y \quad (2)$$

or

$$A_P(x, y, T) = A_P(x, y, Z) \quad \forall x, y$$

This equality guarantees that, if adjusted for, sets  $T$  and  $Z$  would produce the same asymptotic bias relative to the target quantity.

Note that when  $Z$  is a subset of  $T$ ,  $c$ -equivalence amounts to *collapsibility* (of  $T$  over  $T \setminus Z$ ), a topic discussed extensively in biostatistics and epidemiology [14–16] as well as in Section 6 (Theorem 4).

### Definition 2. (Causal admissibility)

Let  $P(y|do(x))$  stand for the “causal effect” of  $X$  on  $Y$ , i.e. the distribution of  $Y$  after setting variable  $X$  to a constant  $X = x$  by external intervention. A set  $Z$  of covariates is said to be “causally admissible” (for adjustment) relative to the causal effect of  $X$  on  $Y$ , if the following equality holds for all  $x \in X$  and all  $y \in Y^2$ :

<sup>1</sup> Integrals should replace summations whenever the variables are continuous.

<sup>2</sup> Equivalently, one can define admissibility using the equality:

$$P(y|do(x)) = P(Y_x = y) \quad (3)$$

where  $Y_x$  is the counterfactual or “potential outcome” variable [17, 18]. The equivalence of the two definitions is shown in Pearl [8].

$$\sum_z P(y|x, z)P(z) = P(y|do(x)) \quad (4)$$

Whereas bias reduction provides a motivation for seeking a set  $Z$  that approximates eq. (4),  $c$ -equivalence, as defined in eq. (2), is not a causal concept, for it depends solely on the properties of the joint probability  $P$ , regardless of the causal connections between  $X, Y, Z$ , and  $T$ . Our aim however is to give a characterization of  $c$ -equivalence, not in terms of a specific distribution  $P(x, y, z)$  but, rather, in terms of qualitative attributes of  $P$  that can be ascertained from scientific knowledge prior to obtaining any data. Since graphs provide a useful and meaningful representation of such knowledge (e.g. in terms of conditional-independence relations) we will aim to characterize  $c$ -equivalence in terms of the graphical relationships among the variables in  $X, Y, Z$ , and  $T$ . This way, the conditions derived will secure  $c$ -equivalence in all distributions  $P$  that share the same graph structure.

To this end, we define the notion of *Markov compatibility*, between a graph  $G$  and a distribution  $P$ .

**Definition 3.** (*Markov compatibility*)

Consider a DAG  $G$  in which each node corresponds to a variable in a probability distribution  $P$ . We say that  $G$  and  $P$  are *Markov compatible* if each variable  $X$  is independent of all its non-descendants, conditioned on its parents in  $G$ . Formally, we write

$$(X \perp\!\!\!\perp nd(X)|pa(X))_P$$

where  $nd(X)$  and  $pa(X)$  are, respectively, the sets of non-descendants and parents of  $X$ .

The set of distributions  $P$  that are compatible with a given DAG  $G$  corresponds to those distributions that can be generated, or simulated by assigning stochastic processors to the arrows in  $G$ , where each processor assigns variable  $X$  a value  $X = x$  according to the conditional probability  $P(X = x|pa(X))$ . Such a process will also be called “parameterization” of  $G$ , since it determines the parameters of the distribution while complying with the structure of  $G$ .

We will say that sets  $T$  and  $Z$  are *c-equivalent* in  $G$ , if they are  $c$ -equivalent in every distribution that is Markov compatible with  $G$ , that is, in every parametrization of  $G$ . However, since  $c$ -equivalence is a probabilistic notion, the causal reading of the arrows in  $G$  can be ignored; what matters is the conditional independencies induced by those arrows, and those are shared by all members of the Markov compatible class. These conditional independencies can be read from  $G$  using a graphical property called “ $d$ -separation.”

**Definition 4.** (*d-separation*)

A set  $S$  of nodes in a graph  $G$  is said to *block a path*  $p$  if either (i)  $p$  contains at least one arrow-emitting node that is in  $S$  or (ii)  $p$  contains at least one collision node that is outside  $S$  and has no descendant in  $S$ . If  $S$  blocks all paths from  $X$  to  $Y$ , it is said to “ $d$ -separate  $X$  and  $Y$ ,” written  $(X \perp\!\!\!\perp Y|S)_G$  and then,  $X$  and  $Y$  are independent given  $S$ , written  $X \perp\!\!\!\perp Y|S$ , in every probability distribution that is compatible with  $G$  [19].

If two DAGs,  $G_1$  and  $G_2$ , induce the same set of  $d$ -separations on a set  $V$  of variables, they are called “Markov equivalent,” and they share the same set of Markov compatible distributions. Clearly, if two sets are  $c$ -equivalent in graph  $G_1$ , they are also  $c$ -equivalent in any graph  $G_2$  that is Markov equivalent to  $G_1$ , regardless of the directionality of their arrows. It is convenient, nevertheless, to invoke the notion of “admissibility” which is causal in nature (see Definition 2), hence sensitive to causal directionality. Admissibility will play a pivotal role in our analysis in Sections 3–5 and will be replaced with a non-causal substitute in Section 6. The next definition casts admissibility in graphical terms and connects it with  $c$ -equivalence.

**Definition 5.** (*G-admissibility*)

Let  $pa(X)$  be the set of  $X$ 's parents in a DAG  $G$ . A set of nodes  $Z$  is said to be *G-admissible* if for every  $P$  compatible with  $G$ ,  $Z$  is  $c$ -equivalent to  $pa(X)$ , namely,

$$\sum_{pa(X)} P(y|x, pa(X))P(pa(X)) = \sum_z P(y|x, z)P(z) \quad \forall x, y \quad (5)$$

Definition 5, however, does not provide a graphical test for admissibility since it relies on the notion of  $c$ -equivalence, for which we seek a graphical criterion. A weak graphical test is provided by the back-door criterion to be defined next:

**Definition 6.** (*The back-door criterion*)

A set  $S$  of nodes in a DAG  $G$  is said to satisfy the “back-door criterion” if the following two conditions hold:

1. No element of  $S$  is a descendant of  $X$
2. The elements of  $S$  “block” all “back-door” paths from  $X$  to  $Y$ , namely all paths that end with an arrow pointing to  $X$ .

Alternatively, Condition 2 can be stated as a  $d$ -separation condition in a modified graph:

$$2'. \quad (X \perp\!\!\!\perp Y | S)_{G_X} \quad (6)$$

where  $G_X$  is the subgraph created by removing all arrows emanating from  $X$ .

**Lemma 1.** A sufficient condition for a set  $Z$  to be  $G$ -admissible (Definition 5) is for  $Z$  to satisfy the back-door criterion (Definition 6).

**Proof:**

Lemma 1 was originally proven in the context of causal graphs [9] where it was shown that the back-door condition leads to causal admissibility (eq. (4)), from which eq. (5) follows. A direct proof of Lemma 1 is given in Pearl [20, p. 133] and is based on the fact that the set of parent  $pa(X)$  is always causally admissible for adjustment.<sup>3</sup> ■

Clearly, if two subsets  $Z$  and  $T$  are  $G$ -admissible, they must be  $c$ -equivalent, for their adjustment estimands coincide with  $A_P(x, y, pa(X))$ , for every  $P$  compatible with  $G$ . Therefore, a trivial graphical condition for  $c$ -equivalence is for  $Z$  and  $T$  to satisfy the back-door criterion of Definition 6. This condition, as we shall see in the next section, is rather weak;  $c$ -equivalence extends beyond admissible sets.

### 3 Conditions for $c$ -equivalence

**Theorem 1.** A sufficient condition for the  $c$ -equivalence of  $T$  and  $Z$  is that  $Z$  satisfies:

$$\begin{aligned} (X \perp\!\!\!\perp Z | T) & \quad (i) \\ (Y \perp\!\!\!\perp T | X, Z) & \quad (ii) \end{aligned}$$

**Proof:**

Conditioning on  $Z$ , (ii) permits us to rewrite the left-hand side of eq. (2) as

$$\begin{aligned} A_P(x, y, T) &= \sum_t P(t) \sum_z P(y|z, x, t)P(z|t, x) \\ &= \sum_t P(t) \sum_z P(y|z, x)P(z|t, x) \end{aligned}$$

<sup>3</sup> When  $G$  is a causal graph,  $A_P(x, y, pa(X))$  coincides with the causal effect  $P(y|do(x))$ , since adjustment for the direct cause,  $pa(X)$ , deconfounds the relationship between  $X$  and  $Y$  [20, p. 74, Theorem 3.2.2]. For proof and intuition behind the back-door test, as well as a relaxation of the requirement of no descendants, see [20, p. 339] and Lemma 4.

Condition (i) further yields  $P(z|t, x) = P(z|t)$ , from which the equality in eq. (2) follows:

$$\begin{aligned} A_P(x, y, T) &= \sum_t \sum_z P(y|z, x)P(z, t) \\ &= \sum_z P(y|z, x)P(z) \\ &= A_P(x, y, Z) \end{aligned}$$

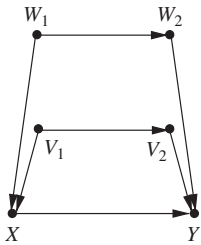
**Corollary 1.** *A sufficient condition for the c-equivalence of T and Z is that either one of the following two conditions holds:*

$$\begin{aligned} C^* &: X \perp\!\!\!\perp Z|T \quad \text{and} \quad Y \perp\!\!\!\perp T|Z, X \\ C^{**} &: X \perp\!\!\!\perp T|Z \quad \text{and} \quad Y \perp\!\!\!\perp Z|T, X \end{aligned} \tag{7}$$

**Proof:**

$C^*$  permits us to derive the right-hand side of eq. (2) from the left-hand side, while  $C^{**}$  permits us to go the other way around. ■

The conditions offered by Theorem 1 and Corollary 1 do not characterize *all* equivalent pairs,  $T$  and  $Z$ . For example, consider the graph in Figure 1, in which each of  $T = \{V_1, W_2\}$  and  $Z = \{V_2, W_1\}$  is  $G$ -admissible they must therefore be  $c$ -equivalent. Yet neither  $C^*$  nor  $C^{**}$  holds in this case.



**Figure 1** The sets  $T = \{V_1, W_1\}$  and  $Z = \{V_2, W_2\}$  satisfy the conditions of Theorem 1. The sets  $T = \{V_1, W_2\}$  and  $Z = \{V_2, W_1\}$  block all back-door paths between  $X$  and  $Y$ , hence they are admissible and  $c$ -equivalent. Still they do not satisfy the conditions of Theorem 1

On the other hand, condition  $C^*$  can detect the  $c$ -equivalence of some non-admissible sets, such as  $T = \{W_1\}$  and  $Z = \{W_1, W_2\}$ . These two sets are non-admissible for they fail to block the back-door path  $X \leftarrow V_1 \rightarrow V_2 \rightarrow Y$ , yet they are  $c$ -equivalent according to Theorem 1; (i) is satisfied by  $d$ -separation, while (ii) is satisfied by subsumption ( $T \subseteq Z$ ).

It is interesting to note however that  $Z = \{W_1, W_2\}$ , while  $c$ -equivalent to  $\{W_1\}$ , is not  $c$ -equivalent to  $T = \{W_2\}$ , though the two sets block the same path in the graph.<sup>4</sup> Indeed, this pair does not meet the test of Theorem 1; choosing  $T = \{W_2\}$  and  $Z = \{W_1, W_2\}$  violates Condition (i) since  $X$  is not  $d$ -separated from  $W_1$ , while choosing  $Z = \{W_2\}$  and  $T = \{W_1, W_2\}$  violates Condition (ii) by unblocking the path  $W_1 \rightarrow X \leftarrow V_1 \rightarrow V_2 \rightarrow Y$ . Likewise, the sets  $T = \{W_1\}$  and  $Z = \{W_2\}$  block the same path and, yet, are not  $c$ -equivalent; they fail indeed to satisfy Condition (ii) of Theorem 1.

We are now ready to broaden the scope of Theorem 1 and derive a condition (Theorem 2) that detects *all*  $c$ -equivalent subsets in a graph, as long as they do not contain descendants of  $X$ .

<sup>4</sup> The reason is that the strength of the association between  $X$  and  $Y$ , conditioned on  $W_2$ , depends on whether we also condition on  $W_1$ . Else,  $P(y|x, w_2)$  would be equal to  $P(y|x, w_1, w_2)$  which would render  $Y$  and  $W_1$  independently given  $X$  and  $W_2$ . But this is true only if the path  $(X, V_1, V_2, Y)$  is blocked. See Pearl [21].

**Definition 7.** (*Markov Blanket*)

For any subset  $S$  of variables of  $G$ , a subset  $S'$  of  $S$  will be called a Markov Blanket (MB) if it satisfies the condition

$$(X \perp\!\!\!\perp S|S')_G \quad (8)$$

**Lemma 2.** Every set of variables,  $S$ , is  $c$ -equivalent to any of its MBs.

**Proof:**

Choosing  $Z = S$  and  $T = S'$  satisfies the two conditions of Theorem 1; (i) is satisfied by the definition of  $S'$  (eq. (8)), while (ii) is satisfied by subsumption ( $T \subseteq Z$ ). ■

**Remark.**

It is shown in Appendix that the set of MBs is closed under union and intersection and that it contains a unique minimal set, denoted  $S_m$ . This leads to the following definition:

**Definition 8.** (*Markov Boundary*)

The unique and minimal MB of a given subset  $S$  with regard to  $X$  will be called the Markov Boundary (MBY) of  $S$  relative to  $X$  (or the MBY of  $S$  when  $X$  is presumed given). Note that the measurement of the MBY renders  $X$  independent of all other members of  $S$  and no other subset of the MBY has this property.

**Lemma 3.** Let  $Z$  and  $T$  be two subsets of vertices of  $G$ . Then  $Z_m = T_m$  if and only if  $(X \perp\!\!\!\perp (Z, T)|S_I)_G$  where  $S_I$  is the intersection of  $Z$  and  $T$ . In words,  $Z$  and  $T$  have identical MBYs iff they are  $d$ -separated from  $X$  by their intersection.

**Proof:**

If the condition holds then  $S_I$  must be a MB of both  $Z$  and  $T$ . So the unique minimal MB of both  $Z$  and  $T$  must be included in  $S_I$  and is the MBY of both sets. If the MBY of both  $Z$  and  $T$  are equal then they must be a subset of  $Z$  and of  $T$  so the condition must hold. ■

**Theorem 2.** Let  $Z$  and  $T$  be two sets of variables in  $G$  containing no descendants of  $X$ . A necessary and sufficient condition for  $Z$  and  $T$  to be  $c$ -equivalent in  $G$  is that at least one of the following two conditions holds:

1.  $(X \perp\!\!\!\perp (Z, T)|S_I)_G$  where  $S_I$  is the intersection of  $Z$  and  $T$
2.  $Z$  and  $T$  are  $G$ -admissible, i.e. they satisfy the back-door criterion.

**Proof:**

Due to Lemma 3 we can replace in our proof Condition 1 by the condition  $Z_m = T_m$ .

1. Proof of sufficiency:

Condition 2 is sufficient since  $G$ -admissibility implies admissibility and renders the two adjustment estimands in eq. (2) equal to the causal effect. Condition 1 is sufficient by reason of Lemma 2, which yields

$$Z \sim Z_m \sim T_m \sim T$$

2. Proof of necessity:

We need to show that if Conditions (1) and (2) are both violated then there is at least one parameterization of  $G^5$  (that is, an assignment of conditional probabilities to the parent-child families in  $G$ ) that violates

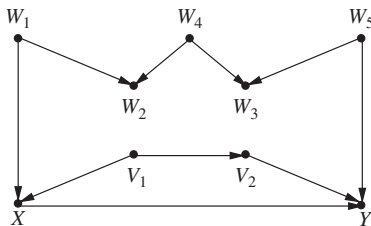
---

<sup>5</sup> In the rest of the paper, we will use the abbreviation  $c$ -equivalent whenever no confusion arises.

eq. (2). If exactly one of  $(Z, T)$  is  $G$ -admissible then  $Z$  and  $T$  are surely not  $c$ -equivalent, for their adjustment estimands would differ for some parameterization of the graph. Assume that both  $Z$  and  $T$  are not  $G$ -admissible or, equivalently, that none of  $Z_m$  or  $T_m$  is  $G$ -admissible. Then there is a back-door path  $p$  from  $X$  to  $Y$  that is not blocked by either  $Z_m$  or  $T_m$ . If, in addition, condition (1) is violated (i.e.  $Z_m$  differs from  $T_m$ ), then  $T_m$  and  $Z_m$  cannot both be disconnected from  $X$  (for then  $Z_m = T_m = \emptyset$ , satisfying condition (1), there must be either a path  $p_1$  from  $Z_m$  to  $X$  that is not blocked by  $T_m$  or a path  $p_2$  from  $T_m$  to  $X$  that is not blocked by  $Z_m$ . Assuming the former case, there must be an unblocked path  $p_1$  from  $Z_m$  to  $X$  followed by a back-door path  $p$  from  $X$  to  $Y$ . The existence of this path implies that, conditional on  $T$  the association between  $X$  and  $Y$  depends on whether we also condition on  $Z$  (see Footnote 4). The fact that the graph permits such dependence means that there exists a parametrization in which such dependence is realized, thus violating the  $c$ -equivalence between  $Z$  and  $T$  (eq. (2)). For example, using a linear parametrization of the graph, we first weaken the links from  $T_m$  to  $X$  to make the left-hand side of eq. (2) equal to  $P(y|x)$ , or  $A(x, y, T_m) = A(x, y, 0)$ . Next, we construct a linear model in which the parameters along paths  $p_1$  (connecting  $Z_m$  to  $X$ ) and the back-door path  $p$  are non-zero. Wooldridge [22] has shown (see also Pearl [21, 23]) that adjustment for  $Z_m$  under such conditions results in a higher bias relative to the unadjusted estimand, or  $A(x, y, Z_m) \neq A(x, y, 0)$ . This completes the proof of necessity, because the parametrization above leads to the inequality  $A(x, y, Z_m) \neq A(x, y, T_m)$ , which implies  $Z \not\sim T$ . ■

### 4 Illustrations

Figure 2 illustrates the power of Theorem 2. In this model, no subset of  $\{W_1, W_2, W_3\}$  is  $G$ -admissible (because of the back-door path through  $V_1$  and  $V_2$ ) and, therefore, equality of MBYs is necessary and sufficient for  $c$ -equivalence among any two such subsets. Accordingly, we can conclude that  $T = \{W_1, W_2\}$  is  $c$ -equivalent to  $Z = \{W_1, W_3\}$ , since  $T_m = W_1$  and  $Z_m = W_1$ . Note that  $W_1$  and  $W_2$ , though they result (upon conditioning) in the same set of unblocked paths between  $X$  and  $Y$ , are not  $c$ -equivalent since  $T_m = W_1 \neq Z_m = W_2$ . Indeed, each of  $W_1$  and  $W_2$  is an instrumental variable relative to  $\{X, Y\}$ , with potentially different strengths, hence potentially different adjustment estimands. Sets  $W_4$  and  $W_3$  however are  $c$ -equivalent, because the MBY of each is the null set,  $\{\emptyset\}$ .



**Figure 2**  $W_3$  and  $W_4$  are non-admissible yet  $c$ -equivalent; both having  $\emptyset$  as a MBY. However,  $W_2$  and  $W_3$  are not  $c$ -equivalent with MBYs  $W_2$  and  $\emptyset$ , respectively

We note that testing for  $c$ -equivalence can be accomplished in polynomial time. The MBY of an arbitrary set  $S$  can be identified by iteratively removing from  $S$ , in any order, any node that is  $d$ -separated from  $X$  given all remaining members of  $S$  (see Appendix 1).  $G$ -admissibility, likewise, can be tested in polynomial time [24].

Theorem 2 also leads to a step-wise process of testing  $c$ -equivalence,

$$T \sim T_1 \sim T_2 \sim \dots \sim Z$$

where each intermediate set is obtained from its predecessor by an addition or deletion of one variable only. This can be seen by organizing the chain into three sections.

$$T \sim \dots \sim T_m \sim \dots \sim Z_m \sim \dots \sim Z$$

The transition from  $T$  to  $T_m$  entails the deletion from  $T$  of all nodes that are not in  $T_m$ ; one at a time, in any order. Similarly, the transition from  $Z_m$  to  $Z$  builds up the full set  $Z$  from its MBY  $Z_m$ ; again, in any order. Finally, the middle section, from  $T_m$  to  $Z_m$ , amounts to traversing a chain of  $G$ -admissible sets, using both deletion and addition of nodes, one at a time. A theorem due to Tian et al. [24] ensures that such a step-wise transition is always possible between any two  $G$ -admissible sets. In case  $T$  or  $Z$  is non-admissible, the middle section must degenerate into an equality  $T_m = Z_m$ , or else,  $c$ -equivalence does not hold.

Figure 2 can be used to illustrate this step-wise transition from  $T = \{W_1, W_2, V_1\}$  to  $Z = \{V_2, W_3\}$ . Starting with  $T$ , we obtain

$$\begin{aligned} T = \{W_1, W_2, V_1\} &\sim \{W_1, V_1\} = T_m \sim \{V_1\} \sim \{V_1, V_2\} \\ &\sim \{V_2\} = Z_m \sim \{V_2, W_3\} = Z \end{aligned}$$

If, however, we were to attempt a step-wise transition between  $T = \{W_1, W_2, V_1\}$  and  $Z = \{W_3\}$ , we would obtain

$$T = \{W_1, W_2, V_1\} \sim \{W_1, V_1\} \sim \{V_1\}$$

and would be unable to proceed toward  $Z_m = \{W_3\}$ . The reason lies in the non-admissibility of  $Z$  which necessitates the equality  $T_m = Z_m$ , contrary to the MBYs shown in the graph.

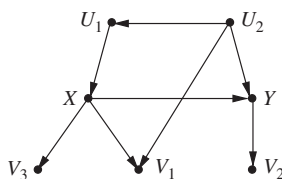
Note also that each step in the process  $T \sim \dots \sim T_m$  (as well as  $Z_m \sim \dots \sim Z$ ) is licensed by Condition (i) of Theorem 1, while each step in the intermediate process  $T_m \sim \dots \sim Z_m$  is licensed by Condition (ii). Both conditions are purely statistical and do not invoke the causal reading of “admissibility.” This means that Condition 2 of Theorem 2 may be replaced by the requirement that  $Z$  and  $T$  satisfy the back-door test in any diagram compatible with  $P(x, y, z, t)$ ; the direction of arrows in the diagram need not convey causal information. Further clarification of the statistical implications of the admissibility condition is given in Section 6.

## 5 Extended conditions for $c$ -equivalence

The two conditions of Theorem 2 are sufficient and necessary as long as we limit the sets  $Z$  and  $T$  to non-descendants of  $X$ . Such sets usually represent “pre-treatment” covariates which are chosen for adjustment in order to reduce confounding bias. In many applications, however, causal-effect estimation is also marred by “selection bias” which occurs when samples are preferentially selected to the data set, depending on the values taken by some variables in the model [23, 25–27]. Selection bias is represented by variables that are permanently conditioned on (to signify selection) and these are often affected by the causal variable  $X$ .

To present a more general condition for  $c$ -equivalence, applicable to any sets of variables, we need to introduce two extensions. First, a graphical criterion for  $G$ -admissibility (Definition 5) must be devised that ensures  $c$ -equivalence with  $pa(X)$  even for sets including descendant of  $X$ . Second, Conditions 1 and 2 in Theorem 2 need to be augmented with a third option, to accommodate new  $c$ -equivalent pairs  $(Z, T)$  that may not meet Conditions 1 and 2.

To illustrate, consider the graph of Figure 3. Clearly, the sets  $\{U_1\}$ ,  $\{U_2\}$ , and  $\{U_1, U_2\}$  all satisfy the back-door criterion and are therefore  $G$ -admissible. The set  $\{V_1\}$  however fails the back-door test on two



**Figure 3** Demonstrating the extended-back-door criterion (Definition 9), which allows admissible sets to include descendants of  $X$ . The sets  $\{V_3, U_1\}$  and  $\{V_1, U_2\}$  are admissible, but not  $\{U_1, V_1\}$  or  $\{U_2, V_2\}$



accounts: it is a descendant of  $X$  and it does not block the back-door path  $X \leftarrow U_1 \leftarrow U_2 \rightarrow Y$ . In addition, conditioning on  $V_1$  opens a non-causal path between  $X$  and  $Y$  which should further disqualify  $V_1$  from admissibility. Consider now the set  $\{V_1, U_2\}$ . This set does block all back-door paths and does not open any spurious (non-causal) path between  $X$  and  $Y$ . We should therefore qualify  $\{V_1, U_2\}$  as  $G$ -admissible. Indeed, we shall soon prove that  $\{V_1, U_2\}$  is  $c$ -equivalent to the other admissible sets in the graphs,  $\{U_1\}$ ,  $\{U_2\}$ , and  $\{U_1, U_2\}$ .

Next consider the set  $S = \{U_1, V_1\}$  which, while blocking the back-door path  $X \leftarrow U_1 \leftarrow U_2 \rightarrow Y$ , also unblocks the collider path  $X \rightarrow V_1 \leftarrow U_2 \rightarrow Y$ . Such sets should not be characterized as  $G$ -admissible, because they are not  $c$ -equivalent to  $pa(X)$ . Conceptually, admissibility requires that, in addition to blocking all back-door paths, conditioning on a set  $S$  should not open new non-causal paths between  $X$  and  $Y$ .

The set  $S = \{U_2, V_2\}$  should be excluded for the same reason, though the spurious path in this case is more subtle;  $V_2$  is a descendant of a virtual collider  $X \rightarrow Y \leftarrow \varepsilon_Y$  where  $\varepsilon_Y$  (not shown explicitly in the graph) represents all exogenous omitted factors in the equation of  $Y$  (See Pearl [20, pp. 339–40]). The next definition, called extended-back-door, provides a graphical criterion for selecting genuinely admissible sets and excluding those that are inadmissible for the reasons explained above. It thus extends the notion of  $G$ -admissibility (Definition 8) to include variables that are descendants of  $X$ .

**Definition 9.** (*Extended-back-door*)

Let a set  $S$  of variables be partitioned into  $S^+ \cup S^-$ , such that  $S^+$  contains all non-descendants of  $X$  and  $S^-$  the descendants of  $X$ .  $S$  is said to meet the extended-back-door criterion if  $S^+$  and  $S^-$  satisfy the following two conditions.

- A.  $S^+$  blocks all back-door paths from  $X$  to  $Y$
- B.  $X$  and  $S^+$  block all paths between  $S^-$  and  $Y$ , namely,  $(S^- \perp\!\!\!\perp Y | X, S^+)_G$ .<sup>6</sup>

**Lemma 4.** Any set meeting the extended-back-door criterion is  $G$ -admissible, i.e. it is  $c$ -equivalent to  $pa(X)$ .

**Proof:**

Since  $S^+$  satisfies the back-door criterion, it is  $c$ -equivalent to  $pa(X)$  by virtue of eq. (5). To show that  $S^+ \sim \{S^+ \cup S^-\}$ , we invoke Theorem 1 with  $T = \{S^+ \cup S^-\}$  and  $Z = S^+$ . Conditions (i) and (ii) of Theorem 1 then translate into:

$$(i) \quad X \perp\!\!\!\perp S^+ | S^+ \cup S^-$$

$$(ii) \quad Y \perp\!\!\!\perp S^+, S^- | X, S^+$$

(i) is satisfied by subsumption, while (ii) follows from Condition B of Definition 9. This proves the equivalence  $S^+ \sim S$  and, since  $S^+ \sim pa(X)$ , we conclude  $S \sim pa(X)$ . ■

The extra  $d$ -separation required in Condition B of Definition 9 offers a succinct graphical test for the virtual-colliders criterion expressed in Pearl [20, pp. 339–40] as well as the “non-causal paths” criterion of Shpitser et al. [28].<sup>7</sup> It forbids any admissible set from containing “improper” descendants of  $X$ , that is, intermediate nodes on the causal path from  $X$  to  $Y$  as well as any descendants of such nodes. In Figure 5, for example, Lemma 4 concludes that the sets  $\{U_2, V_3\}$  and  $\{U_2, V_1\}$  are both  $G$ -admissible and therefore  $c$ -equivalent.

<sup>6</sup> In causal analysis, Condition B ensures that  $S^-$  does not open any spurious (i.e. non-causal) path between  $X$  and  $Y$ . For example, it excludes from  $S^-$  all nodes that intercept causal paths from  $X$  to  $Y$  as well as descendants of such nodes. See Pearl [20, p. 399] and Shpitser et al. [28] for intuition and justification.

<sup>7</sup> This condition can be viewed as a consequence of Theorem 7 of Shpitser et al. [28], with  $L = \{\emptyset\}$ . However, here the  $d$ -separation is applied to the original graph and the exclusion of “improper” descendants of  $X$  is not imposed a priori. Rather it follows from Theorem 1 and the requirement of  $G$ -admissibility as expressed in eq. (5).

The  $G$ -admissibility of  $\{U_2, V_3\}$  is established by the condition  $\{V_3 \perp\!\!\!\perp Y|X, U_2\}_G$ , whereas that of  $\{U_2, V_3\}$  by  $\{V_1 \perp\!\!\!\perp Y|X, U_2\}_G$ . On the other hand, the sets  $\{U_1, V_1\}$  and  $\{U_1, V_2\}$  are not  $G$ -admissible. The former because it opens a non-causal path  $X \rightarrow V_1 \leftarrow U_2 \rightarrow Y$  between  $X$  and  $Y$  and the latter because  $V_2$  is a descendant of  $Y$  and thus it opens a virtual collider at  $Y$ . Indeed, the set  $S = \{V_2, U_2\}$  violates Condition  $B$  of Lemma 4, since  $X$  and  $S^+ = \{U_2\}$  do not block all paths from  $Y$  to  $S^- = \{V_2\}$ .

We are ready now to characterize sets that violate the two conditions of Theorem 2 and still, by virtue of containing descendants of  $X$  are nevertheless  $c$ -equivalent. Consider the sets  $Z = \{U_2, V_1, V_2\}$  and  $T = \{U_2, V_3, V_2\}$  in Figure 3. Due to the inclusion of  $V_2$ ,  $Z$  and  $T$  are clearly inadmissible. Likewise, their MBYs are, respectively,  $Z_m = Z$  and  $T_m = T$ , which are not identical. Thus,  $Z$  and  $T$  violate the two conditions of Theorem 2, even allowing for the extended version of the back-door criterion. They are nevertheless  $c$ -equivalent as can be seen from the fact that both are  $c$ -equivalent to their intersection  $S_I = \{U_2, V_2\}$ , since  $\{S_I \cup X\}$   $d$ -separates both  $Z$  and  $T$  from  $Y$ , thus complying with the requirements of Theorem 1.

The following lemma generalizes this observation formally.

**Lemma 5.** *Let  $Z$  and  $T$  be any two sets of variables in a graph  $G$  and  $S_I$  their intersection. A sufficient condition for  $Z$  and  $T$  to be  $c$ -equivalent is that  $\{Z \cup T\}$  is  $d$ -separated from  $Y$  by  $\{X \cup S_I\}$ , that is,  $(Y \perp\!\!\!\perp (Z, T)|X, S_I)_G$ .*

**Proof:**

We will prove Lemma 5 by showing that  $T$  (similarly  $Z$ ) is  $c$ -equivalent to  $S_I$ . Indeed, substituting  $S_I$  for  $Z$  in Theorem 1 satisfies Conditions (i) and (ii); the former by subsumption, the latter by the condition  $(Y \perp\!\!\!\perp (Z, T)|X, S_I)_G$  of Lemma 5. ■

**Remark.**

When  $Z$  and  $T$  contain only non-descendants of  $X$ , Lemma 5 implies at least one of the conditions of Theorem 2.

**Theorem 3.** *Let  $Z$  and  $T$  be any two sets of variables in a graph  $G$ . A sufficient condition for  $Z$  and  $T$  to be  $c$ -equivalent is that at least one of the following three conditions holds:*

1.  $(X \perp\!\!\!\perp (Z, T)|S_I)_G$  where  $S_I$  is the intersection of  $Z$  and  $T$
2.  $Z$  and  $T$  are  $G$ -admissible
3.  $(Y \perp\!\!\!\perp (Z, T)|S_I, X)_G$  where  $S_I$  is the intersection of  $Z$  and  $T$

**Proof:**

That Condition 3 is sufficient for  $Z \sim T$  is established in Lemma 5. The sufficiency of Condition 2 stems from the fact that  $G$ -admissibility implies  $Z \sim pa(X) \sim T$ . It remains to demonstrate the sufficiency of Condition 1, but this is proven in Lemmas 2 and 3 which are not restricted to non-descendants of  $X$ . We conjecture that Conditions 1–3 are also necessary. ■

Theorem 3 reveals non-trivial patterns of  $c$ -equivalence that emerge through the presence of non-descendants of  $X$ . It shows for example a marked asymmetry between confounding bias and selection bias. In the former, illustrated in Figure 1, it was equality of the MBYs around  $X$  that ensures  $c$ -equivalence (e.g.  $W_1 \sim \{W_1 W_2\}$  in Figure 1). In the case of selection bias, on the other hand, it is equality of the MBYs around  $Y$  (augmented by  $X$ ) that is required to ensure  $c$ -equivalence. In Figure 3, for example, the  $c$ -equivalence  $\{V_2, U_2\} \sim \{V_2, U_2, U_1\}$  is sustained by virtue of the equality of the MBYs around  $Y$ ,  $\{V_2, U_2, X\}$ . The sets  $\{V_2, U_1\}$  and  $\{V_2, U_2, U_1\}$ , on the other hand, are not equivalent, though they share MBYs around  $X$ .

Another implication of Theorem 3 is that, in the absence of confounding bias, selection bias is invariant to conditioning on instruments. For example, if we remove the arrow  $U_2 \rightarrow Y$  in Figure 3,  $U_1$  and  $U_2$  would then represent two instruments of different strengths (relative to  $X \rightarrow Y$ ). Still, the two have no effect on the selection bias created by conditioning on  $V_2$ , since the sets  $\{V_2, U_1\}$ ,  $\{V_2, U_2\}$ , and  $\{V_2\}$  are  $c$ -equivalent.

## 6 From causal to statistical characterization

Theorem 2, while providing a necessary and sufficient condition for  $c$ -equivalence, raises an interesting theoretical question. Admissibility is a causal notion (i.e. resting on causal assumptions about the direction of the arrows in the diagram, or the identity of  $pa(X)$ , Definition 6) while  $c$ -equivalence is purely statistical. Why need one resort to causal assumptions to characterize a property that relies on no such assumption? Evidently, the notion of  $G$ -admissibility as it was used in the proof of Theorem 2 was merely a surrogate carrier of statistical information; its causal reading, especially the identity of the parent set  $pa(X)$  (Definition 6) was irrelevant. The question then is whether Theorem 2 could be articulated using purely statistical conditions, avoiding admissibility altogether, as is done in Theorem 1.

We will show that the answer is positive; Theorem 2 can be rephrased using a statistical test for  $c$ -equivalence. It should be noted though, that the quest for statistical characterization is of merely theoretical interest; rarely is one in possession of prior information about conditional independencies (as required by Theorem 1), that is not resting on causal knowledge (of the kind required by Theorem 2). The utility of statistical characterization surfaces when we wish to confirm or reject the structure of the diagram. We will see that the statistical reading of Theorem 2 has testable implication that, if failed to fit the data, may help one select among competing graph structures.

Our plan is, first, to obtain a statistical  $c$ -equivalence test for the special case where  $T$  is a subset of  $Z$ , then extend it to arbitrary sets,  $T$  and  $Z$ .

**Theorem 4.** (*Set-subset equivalence – collapsibility*)

Let  $T$  and  $S$  be two disjoint sets. A sufficient condition for the  $c$ -equivalence of  $T$  and  $Z = T \cup S$  is that  $S$  can be partitioned into two subsets,  $S_1$  and  $S_2$ , such that:

$$(i') \quad S_1 \perp\!\!\!\perp X|T$$

and

$$(ii') \quad S_2 \perp\!\!\!\perp Y|S_1, X, T$$

**Proof:**

Starting with

$$A(x, y, T \cup S) = \sum_t \sum_{s_1} \sum_{s_2} P(y|x, t, s_1, s_2)P(s_1, s_2, t)$$

(ii') permits us to remove  $s_2$  from the first factor and write

$$\begin{aligned} A(x, y, T \cup S) &= \sum_t \sum_{s_1} \sum_{s_2} P(y|x, t, s_1)P(s_1, s_2, t) \\ &= \sum_t \sum_{s_1} P(y|x, t, s_1)P(s_1, t) \end{aligned}$$

while (i') permits us to reach the same expression from  $A(x, y, T)$ :

$$\begin{aligned} A(x, y, T) &= \sum_t \sum_{s_1} P(y|x, t, s_1)P(s_1|x, t)P(t) \\ &= \sum_t \sum_{s_1} P(y|x, t, s_1)P(s_1, t) \end{aligned}$$

which proves the theorem. ■

Theorem 4 can also be proven by double application of Theorem 1; first showing the  $c$ -equivalence of  $T$  and  $\{T \cup S_1\}$  using (i) (with (ii) satisfied by subsumption), then showing the  $c$ -equivalence of  $\{T \cup S_1\}$  and  $\{T \cup S_1 \cup S_2\}$  using (ii) (with (i) satisfied by subsumption).

The advantage of Theorem 4 over Theorem 1 is that it allows certain cases of  $c$ -equivalence to be verified in a single step. In Figure 1, for example, both (i') and (i'') are satisfied for  $T = \{V_1, W_2\}$ ,  $S_1 = \{V_2\}$ , and  $S_2 = \{W_1\}$ . Therefore,  $T = \{V_1, W_2\}$  is  $c$ -equivalent to  $\{T \cup S\} = \{V_1, V_2, W_1, W_2\}$ . While this equivalence can be established using Theorem 1, it would have taken us two steps: first  $T = \{V_1, W_2\} \sim \{V_1, W_2, W_1\}$ , and then  $\{V_1, W_2, W_1\} \sim \{V_1, W_2, W_1, V_2\} = \{T \cup S\}$ .

Theorem 4 in itself does not provide an effective way of testing the existence of a partition  $S = S_1 + S_2$ . However, Appendix 1 shows that a partition satisfying the conditions of Theorem 4 exists if and only if  $S_2$  is the (unique) maximal subset of  $S$  that satisfies

$$\{Y \perp\!\!\!\perp S_2 | (S - S_2), X, T\}$$

In other words,  $S_2$  can be constructed incrementally by selecting each and only elements  $s_i$  satisfying

$$\{Y \perp\!\!\!\perp s_i | (S - s_i), X, T\}.$$

This provides a linear algorithm for testing the existence of a desired partition and, hence, the  $c$ -equivalence of  $T$  and  $S + T$ .

Theorem 4 generalizes closely related theorems by Stone [7] and Robins [29], in which  $T \cup S$  is assumed to be admissible (see also Greenland et al. [30]). The importance of this generalization was demonstrated by several examples in Section 3. Theorem 4 on the other hand invokes only the distribution  $P(x, y, z, t)$  and makes no reference to  $P(y|do(x))$  or to admissibility.

The weakness of Theorem 4 is that it is applicable to set-subset relations only. A natural attempt to generalize the theorem would be to posit the requirement that  $T$  and  $Z$  each be  $c$ -equivalent to  $T \cup Z$  and use Theorem 4 to establish the required set-subset equivalence. While perfectly valid, this condition is still not complete; there are cases where  $T$  and  $Z$  are  $c$ -equivalent, yet none is  $c$ -equivalent to their union. For example, consider the path

$$X \rightarrow T \leftarrow L \rightarrow Z \leftarrow Y$$

Each of  $T$  and  $Z$  leaves the path between  $X$  and  $Y$  blocked, which renders them  $c$ -equivalent, yet  $\{T \cup Z\}$  unblocks that path. Hence,  $T \sim Z$  and  $T \not\sim T \cup Z$ . This implies that sets  $T$  and  $Z$  would fail the proposed test, even though they are  $c$ -equivalent.

The remedy can be obtained by re-invoking the notion of MBY (Definition 8) and Lemma 2.

**Theorem 5.** *Let  $T$  and  $Z$  be two sets of covariates, containing no descendant of  $X$  and let  $T_m$  and  $Z_m$  be their MBYs. A necessary and sufficient condition for the  $c$ -equivalence of  $T$  and  $Z$  is that each of  $T_m$  and  $Z_m$  be  $c$ -equivalent to  $T_m \cup Z_m$  according to the set-subset criterion of Theorem 4.*

**Proof:**

1. Proof of sufficiency:

If  $T_m$  and  $Z_m$  are each  $c$ -equivalent to  $T_m \cup Z_m$ , then, obviously, they are  $c$ -equivalent themselves and, since each is  $c$ -equivalent to its parent set (by Lemma 2)  $T$  and  $Z$  are  $c$ -equivalent as well.

2. Proof of necessity:

We need to show that if either  $T_m$  or  $Z_m$  is not  $c$ -equivalent to their union (by the test of Theorem 4), then they are not  $c$ -equivalent to each other. We will show that using “ $G$ -admissibility” as an auxiliary tool. We will show that failure of  $Z_m \sim T_m \cup Z_m$  implies non-admissibility, and this, by the necessary part of Theorem 2, negates the possibility of  $c$ -equivalence between  $Z$  and  $T$ . The proof relies on the monotonicity of  $d$ -separation over minimal subsets (Appendix 2), which states that, for any graph  $G$ , and any two subsets of nodes  $T$  and  $Z$ , we have

$$(X \perp\!\!\!\perp Y | Z_m)_G \ \& \ (X \perp\!\!\!\perp Y | T_m)_G \ \Rightarrow \ (X \perp\!\!\!\perp Y | Z_m \cup T_m)_G$$

Applying this to the subgraph consisting of all back-door paths from  $X$  to  $Y$ , we conclude that  $G$ -admissibility is preserved under union of minimal sets. Therefore, the admissibility of  $Z_m$  and  $T_m$  (hence of  $Z$  and  $T$ ) entails admissibility of  $Z_m \cup T_m$ . Applying Theorem 2, this implies the necessity part of Theorem 4. ■

Theorem 5 reveals the statistical implications of the  $G$ -admissibility requirement in Theorem 2.  $G$ -admissibility ensures the two  $c$ -equivalence conditions:

$$T_m \sim \{T_m \cup Z_m\} \quad (9)$$

$$Z_m \sim \{T_m \cup Z_m\} \quad (10)$$

In other words, given any DAG  $G$  compatible with the conditional independencies of  $P(x, y, t, z)$ , whenever  $Z$  and  $T$  are  $G$ -admissible in  $G$ , the two statistical conditions of Theorem 4 should hold in the distribution and satisfy the equivalence relationships in eqs (9) and (10). Explicating these two conditions using the proper choices of  $S_1$  and  $S_2$  yields

$$\{T_m \cup Z_m\}_m \setminus Z_m \perp\!\!\!\perp Y|X, Z_m, T_m \setminus \{T_m \cup Z_m\}_m \quad (11)$$

$$\{T_m \cup Z_m\}_m \setminus T_m \perp\!\!\!\perp Y|X, T_m, Z_m \setminus \{T_m \cup Z_m\}_m \quad (12)$$

which constitute the statistical implications of admissibility. These implications should be confirmed in any graph  $G'$  that is Markov equivalent to  $G$ , regardless of whether  $T$  and  $S$  are  $G$ -admissible in  $G'$  and regardless of the identity of  $pa(X)$  in  $G'$ .

We illustrate these implications using Figure 2. Taking  $T = \{W_2, V_2\}$  and  $Z = \{V_1, W_3\}$ , we have

$$\begin{aligned} T_m &= \{W_2, V_2\}, Z_m = \{V_1\}, \\ \{T_m \cup Z_m\}_m &= \{V_1, V_2, W_2\}_m = \{V_1, W_2\} \end{aligned}$$

We find that the tests of eqs (11) and (12) are satisfied because

$$\{W_2\} \perp\!\!\!\perp Y|X, V_1, V_2 \quad \text{and} \quad V_1 \perp\!\!\!\perp Y|X, W_2, V_2$$

Thus, implying  $Z \sim T$ . That test would fail had we taken  $T = \{W_2\}$  and  $Z = \{W_3\}$ , because then we would have

$$\begin{aligned} T_m &= \{W_2\}, Z_m = \{\emptyset\}. \\ \{T_m \cup Z_m\}_m &= W_2 \end{aligned}$$

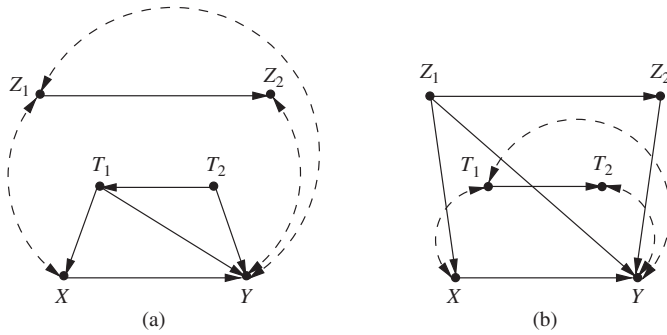
and the requirement

$$\{T_m \cup Z_m\}_m \setminus Z_m \perp\!\!\!\perp Y|X, Z_m, T_m \setminus \{T_m \cup Z_m\}_m$$

would not be satisfied because

$$W_2 \not\perp\!\!\!\perp Y|X$$

Figure 4 presents two models that are observationally indistinguishable, yet they differ in admissibility claims. Model 4(a) deems  $\{T_1\}$  and  $\{T_1, T_2\}$  to be admissible, while Model 4(b) counters (a) and deems  $\{Z_1\}$  and  $\{Z_1, Z_2\}$  to be admissible. Indistinguishability requires that  $c$ -equivalence be preserved and, indeed, the relations  $\{T_1\} \approx \{T_1, T_2\}$  and  $\{Z_1\} \approx \{Z_1, Z_2\}$  are held in both (a) and (b).



**Figure 4** Two observationally indistinguishable models that differ in their admissible sets. Both confirm the  $c$ -equivalence  $\{T_1\} \sim \{T_1, T_2\}$  and  $\{Z_1\} \sim \{Z_1, Z_2\}$  but for different reasons

## 7 Empirical ramifications of $c$ -equivalence tests

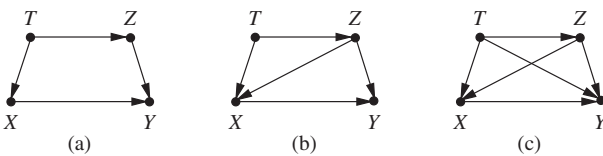
Having explicated the statistical implications of admissibility vis-a-vis  $c$ -equivalence, we may ask the inverse question: What can  $c$ -equivalence tests tell us about admissibility? It is well known that no statistical test can ever confirm or refute the admissibility of a given set  $Z$  (Pearl [8], Chapter 6, Pearl [31]). The discussion of Section 6 shows however that the admissibility of two sets,  $T$  and  $Z$ , does have testable implications. In particular, if they fail the  $c$ -equivalence test, they cannot both be admissible. This might sound obvious, given that admissibility entails zero bias for each of  $T$  and  $Z$  (eq. (7)). Still, eq. (10) implies that it is enough for  $Z_m$  (or  $T_m$ ) to fail the  $c$ -equivalence test vis-a-vis  $\{Z_m \cup T_m\}$  for us to conclude that, in addition to having different MBYs,  $Z$  and  $T$  cannot both be admissible.

This finding can be useful when measurements need be chosen (for adjustment) with only partial knowledge of the causal graph underlying the problem. Assume that two candidate graphs recommend two different measurements for confounding control, one graph predicts the admissibility of  $T$  and  $Z$ , and the second does not. Failure of the  $c$ -equivalence test

$$T_m \sim \{T_m \cup Z_m\} \sim Z_m$$

can then be used to rule out the former.

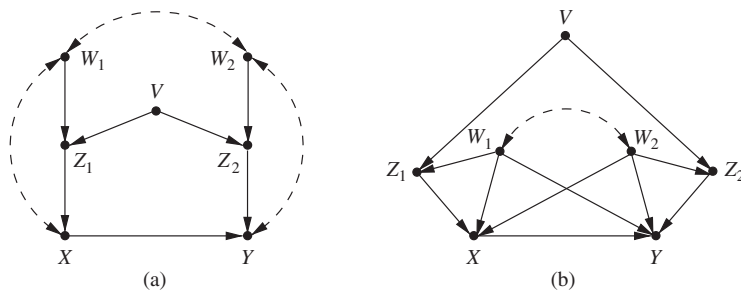
Figure 5 illustrates this possibility. Model 5(a) deems measurements  $T$  and  $Z$  as equally effective for bias removal, while models 5(b) and 5(c) deem  $T$  to be insufficient for adjustment. Submitting the data to the  $c$ -equivalence tests of eqs. (9) and (10) may reveal which of the three models should be ruled out. If both tests fail, we must rule out Models 5(a) and 5(b), while if only eq. (10) fails, we can rule out only Model 2(a) (eq. (9) may still be satisfied in Model 5(c) by incidental cancellation). This is an elaboration of the “Change-in-Estimate” procedure used in epidemiology for confounder identification and selection [32]. Evans et al. [33] used similar considerations to select and reject DAGs by comparing differences among effect estimates of several adjustment sets against the differences implied by the DAGs.



**Figure 5** Failing the  $T \sim \{T \cup Z\}$  test should reject Model (a) in favor of (b) or (c). Failing  $Z \sim \{T \cup Z\}$  should reject Models (a) and (b) in favor of (c)

Of course, the same model exclusion can be deduced from conditional-independence tests. For example, Models 5(a) and 5(b) both predict  $T \perp\!\!\!\perp Y|X, Z$  which, if violated in the data, would leave Model 5(c) as our choice and behoove us to adjust for both  $T$  and  $Z$ . However, when the dimensionality of the conditioning sets increases, conditional-independence tests are both unreliable and computationally expensive. Although both  $c$ -equivalent and conditional-independence tests can reap the benefits of propensity scores methods (see Appendix 3) which reduce the dimensionality of the conditioning set to a single scalar, it is not clear where the benefit can best be realized, since the cardinalities of the sets involved in these two types of tests may be substantially different.

Figure 6 illustrates this potential more acutely. It is not easy to tell whether Models (a) and (b) are observationally distinguishable, since they embody the same set of missing edges. Yet whereas Model 6(a) has no admissible set (among the observables), its contender, Model 6(b) has three (irreducible) such sets:  $\{Z_1, W_1, W_2\}$ ,  $\{W_1, W_2, Z_2\}$ , and  $\{V, W_1, W_2\}$ . This difference in itself does not make the two models distinguishable (see Figure 4); for example,  $X \rightarrow Z \rightarrow Y$  is indistinguishable from  $X \leftarrow Z \rightarrow Y$ , yet  $Z$  is admissible in the latter, not in the former. However, noting that the three admissible subsets of 6(b) are not  $c$ -equivalent in 6(a) – their MBYs differ – tells us immediately that the two models differ in their statistical implications. Indeed, Model 6(b) should be rejected if any pair of the three sets fails the  $c$ -equivalence test.



**Figure 6** The model in (b) is almost indistinguishable from that of (a), save for advertising one additional independency:  $\{Z_1, V\} \perp\!\!\!\perp Y|X, W_1, W_2, Z_2$ . It deems three sets to be admissible (hence  $c$ -equivalent):  $\{V, W_1, W_2\}$ ,  $\{Z_1, W_1, W_2\}$ , and  $\{W_1, W_2, Z_2\}$ , and would be rejected therefore if any pair of them fails the  $c$ -equivalence test. No such pair is deemed  $c$ -equivalent in Model 6 (a) where the three MBYs are distinct

Visually, the statistical property that distinguishes between the two models is not easy to identify. If we list systematically all their conditional-independence claims, we find that both models share the following:

$$\begin{array}{ll}
 V \perp\!\!\!\perp \{W_1, W_2\} & X \perp\!\!\!\perp \{V, Z_2\}|\{Z_1, W_2, W_1\} \\
 Z_1 \perp\!\!\!\perp \{W_2, Z_2\}|\{V, W_2\} & V \perp\!\!\!\perp Y|\{X, Z_2, W_2, Z_1, W_1\} \\
 Z_2 \perp\!\!\!\perp \{W_1, Z_1, X\}|\{V, W_2\} & V \perp\!\!\!\perp Y|\{Z_2, W_2, Z_1, W_1\}
 \end{array}$$

They disagree however on one additional (and obscured) independence relation,  $Z_1 \perp\!\!\!\perp Y|X, W_1, W_2, Z_2, V$ , that is embodied in Model 6(b) and not in 6(a). The pair  $(Z_1, Y)$ , though non-adjacent, has no separating set in the diagram of Figure 6(a). While a search for such distinguishing independency can be tedious,  $c$ -equivalence comparisons tell us immediately where models differ and how their distinguishing characteristic can be put to a test.

This raises the interesting question of whether the discrimination power of  $c$ -equivalence equals that of conditional-independence tests. We know from Theorem 5 that all  $c$ -equivalence conditions can be derived from conditional-independence relations. The converse, however, is an open question if we allow  $(X, Y)$  to vary over all variable pairs.

## 8 Conclusions

Theorem 2 provides a simple graphical test for deciding whether one set of pre-treatment covariates has the same bias-reducing potential as another. The test requires either that both sets satisfy the back-door criterion or that  $X$  be  $d$ -separated from the two sets, conditioned on their intersection. Both conditions can be tested by fast, polynomial time algorithms, and could be used to guide researchers in deciding what measurements are worth taking, considering differences in costs, dimensionality, accuracy, and sampling variability.

Theorem 3 extends these results to include post-treatment variables by, first, generalizing the back-door criterion to permit post-treatment variables and, second, providing three  $d$ -separation conditions, either one of which ensures  $c$ -equivalence. We have further shown that the conditions above can be given purely associational interpretation, without invoking notions such as “back-door” or “admissibility” which, in themselves, cannot be defined by associations alone (see Pearl [8], Chapter 6, Pearl [31]).

Finally, we show that that  $c$ -equivalence tests could serve as valuable tools for model selection, and we postulate that such tests can be used in a systematic search for graph structures that are compatible with the data.

**Acknowledgments:** This research was supported in parts by grants from NIH #1R01 LM009961-01, NSF #IIS-0914211, #IIS-1249822, and #IIS-1302448, and ONR #N000-14-09-1-0665, #N0014-13-1-0153, and #N00014-10-1-0933.

## Appendix 1

In this Appendix, we prove a theorem that provides a linear-time test for the conditions of Theorem 4. The proof is based on the five graphoid axioms [19, 34] and is valid therefore for all strictly positive distributions. In particular it is valid for dependencies represented in DAGs.

**Theorem 6.** *Let  $Q, R, S$  be disjoint subsets of variables and let  $P = \{(S_i, S'_i) : S_i S'_i = S\}$  be the set of all partitions of  $S$  that satisfy the following relation:*

$$(Q \perp\!\!\!\perp S_i | S'_i R). \quad (13)$$

*Then*

- a. The left sets  $S_i$  and the right sets  $S'_i$  of the partitions in  $P$  are closed under union and intersection*
- b. The left sets  $S_i$  of the partitions in  $P$  are also closed under subsets, i.e. if  $(S_i, S'_i)$  satisfies eq. (13), then any other partition  $(S_j, S'_j)$  such that  $S_j$  is a subset of  $S_i$ , also satisfies eq. (13).*

**Proof:** Assume that  $(S_i, S'_i)$  and  $(S_j, S'_j)$  are in  $P$ . Split  $S$  into four disjoint subsets  $S = S_1 S_2 S_3 S_4$  such that  $S_i = S_1 S_2, S'_i = S_3 S_4, S_j = S_1 S_3, S'_j = S_2 S_4$ . It follows from the assumption that

$$(Q \perp\!\!\!\perp S_1 S_2 | S_3 S_4 R) \quad \text{and} \quad (Q \perp\!\!\!\perp S_1 S_3 | S_2 S_4 R) \quad (14)$$

By decomposition we get from eq. (14)

$$(Q \perp\!\!\!\perp S_2 | S_3 S_4 R) \quad \text{and} \quad (Q \perp\!\!\!\perp S_3 | S_2 S_4 R) \quad (15)$$

From eq. (15) we get by intersection

$$(Q \perp\!\!\!\perp S_2 S_3 | S_4 R) \quad (16)$$

From eq. (14) we get by weak union

$$(Q \perp\!\!\!\perp S_1 | S_2 S_3 S_4 R) \quad (17)$$



Finally we get from eqs. (16) and (17) by contraction

$$(Q \perp\!\!\!\perp S_1 S_2 S_3 | S_4 R) \quad (18)$$

Property *a* now follows from eqs. (17) and (18), since  $S_1$  is the intersection of  $S_i$  and  $S_j$  and  $S_2 S_3 S_4$  is the union of  $S'_i$  and  $S'_j$ . Similarly  $S_1 S_2 S_3$  is the union of  $S_i$  and  $S_j$  and  $S_4$  is the intersection of  $S'_i$  and  $S'_j$ . Property *b* follows, by weak union from eq. (14) since  $(Q \perp\!\!\!\perp S_i | S'_i, R)$  implies  $(Q \perp\!\!\!\perp S_j | S'_j R)$  when  $S_j$  is a subset of  $S_i$ .

**Corollary 2.** *There is a unique partition in  $P$ ,  $(S_{min}, S'_{max})$  and a unique partition in  $P$ ,  $(S_{max}, S'_{min})$ .*

This follows from property *a*.

**Corollary 3.**

$$S_{max} = \{s_i \text{ in } S : s_i \text{ satisfies } (Q \perp\!\!\!\perp s_i | (S - s_i) R)\}, S'_{min} = S - S_{max} \quad (19)$$

**Proof:** All nodes in eq. (19) satisfy eq. (13) and therefore, by property *a*, their union satisfies eq. (13). On the other hand, any node in  $S_{max}$  must satisfy eq. (19) by property *b*. ■

**Remark.**

If and only if the set  $P$  is not empty then  $S_{max}$  is not empty. This follows from property *b*.

## An algorithm for verifying the conditions of Theorem 4

A simple linear algorithm based on Appendix 1 (where  $Q$  is reset to  $Y$  and  $R$  is reset to  $XT$ ) for verifying the conditions of Theorem 4 is given as follows.

1. Let  $S_2$  be the set of all variables  $s_i$  in  $S$  satisfying the relation

$$(Y \perp\!\!\!\perp s_i | (S - s_i), X, T) \text{ and set: } S_1 = S - S_2.$$

Then  $S_2 = S_{max}$  and  $S_1 = S_{min}$ .

2. There exists a partition  $(S', S'')$  satisfying the conditions of Theorem 4 if and only if  $S_2$  as defined above is not empty and  $S_1$  as defined above satisfies the Condition (*i'*) of Theorem 4.

**Proof:**

IF:  $S_{max}$  satisfies the (*ii'*) condition by its definition. Therefore if  $S_{min}$  satisfies (*i'*) then  $(S_{min}, S_{max})$  is a partition as required given that  $S_{max}$  is not empty.

ONLY IF. If a partition as required  $(S', S'')$  exists, then necessarily  $S_{min}$  is a subset of  $S'$ . Therefore, given that  $S'$  satisfies (*i'*),  $S_{min}$  satisfies this condition too, by decomposition. ■

Notice that if  $S_2$  is empty, then the set of partitions that satisfy the condition (*ii'*) of the theorem is empty by the observation at the end of the Appendix.

## Appendix 2

We prove that, for any graph  $G$ , and any two subsets of nodes  $T$  and  $Z$ , we have

$$(X \perp\!\!\!\perp Y | Z_m)_G \ \& \ (X \perp\!\!\!\perp Y | T_m)_G \Rightarrow (X \perp\!\!\!\perp Y | Z_m \cup T_m)_G$$

where  $Z_m$  and  $T_m$  are any minimal subsets of  $Z$  and  $T$  that satisfy  $(X \perp\!\!\!\perp Y | Z_m)_G$  and  $(X \perp\!\!\!\perp Y | T_m)_G$ , respectively.

The following notation will be used in the proof: A TRAIL will be a sequence of nodes  $v_1, \dots, v_k$  such that  $v_i$  is connected by an arc to  $v_{i+1}$ . A collider  $Z$  is EMBEDDED in a trail if two of his parents belong to the trail. A PATH is a trail that has no embedded collider. We will use the “moralized graph” test of Lauritzen et al. [35] to test for  $d$ -separation (“ $L$ -test,” for short).

**Theorem 7.** *Given a DAG and two vertices  $x$  and  $y$  in the DAG and a set  $\{Z_1, \dots, Z_k\}$  of minimal separators between  $x$  and  $y$ . The union of the separators in the set, denoted by  $Z!$ , is a separator.*

**Proof:**

We mention first two observations:

- (a) Given a minimal separator  $Z$  between  $x$  and  $y$ . If  $Z$  contains a collider  $w$  then there must be a path between  $x$  and  $y$  which is intercepted by  $w$ , implying that  $w$  is an ancestor of either  $x$  or  $y$  or both. This follows from the minimality of  $Z$ . If the condition does not hold, then  $w$  is not required in  $Z$ .
- (b) It follows from (a) above that  $w$  as defined in (a) and its ancestors must belong to the ancestral subgraph of  $x$  and  $y$ .

Let us apply the  $L$ -test to the triplet  $(x, y|Z_1)$ . As  $Z_1$  is a separator, the  $L$ -test must show this. In the first stage of the  $L$ -test, the ancestral graph of the above triplet is constructed. By observation (b) it must include all the colliders that are included in any  $Z_i$ . In the next stage of the  $L$ -test, the parents of all colliders in the ancestral graph are moralized and the directions removed. The result will be an undirected graph including all the colliders in the separators  $Z_i$  and their moralized parents and their ancestors. In this resulting graph,  $Z_1$  still separates between  $x$  and  $y$ . Therefore adding to  $Z_1$  all the colliders in  $Z_i$ ,  $i = 1$  to  $k$ , will result in a larger separator. Adding the noncolliders from all the  $Z_i$  to  $Z_1$  will still keep the separator property of the enlarged set of vertices (trivial). It follows that  $Z!$  is a separator. ■

## Appendix 3

Let the propensity score  $L(z)$  stand for  $P(X = 1|z)$ . It is well known [11] that, viewed as a random variable,  $L(z)$  satisfies  $X \perp\!\!\!\perp L(z)|Z$ . This implies that  $A(x, y, L(z)) = A(x, y, Z)$  and, therefore, testing for the  $c$ -equivalence of  $Z$  and  $T$  can be reduced to testing the  $c$ -equivalence of  $L(z)$  and  $L(t)$ . The latter offers the advantage of dimensionality reduction, since  $L(z)$  and  $L(t)$  are scalars, between zero and one (see Pearl [20, pp. 348–52]).

The same advantage can be utilized in testing conditional independence. To test whether  $(X \perp\!\!\!\perp Y|Z)$  holds in a distribution  $P$ , it is necessary that  $(X \perp\!\!\!\perp Y|L(z))$  holds in  $P$ . This follows from the Contraction axiom of conditional independence, together with the fact that  $Z$  subsumes  $L$ . Indeed, the latter implies

$$X \perp\!\!\!\perp Y|Z \Leftrightarrow X \perp\!\!\!\perp Y|L(z), Z$$

which together with  $X \perp\!\!\!\perp L(z)|Z$  gives

$$X \perp\!\!\!\perp Z|L(z) \ \& \ X \perp\!\!\!\perp Y|L(z), Z \Rightarrow X \perp\!\!\!\perp Y|L(z)$$

The converse requires an assumption of faithfulness.

## References

1. Dawid A. [Influence diagrams for causal modelling and inference](#). Int Stat Rev 2002;70:161–89.
2. Glymour M, Greenland S. Causal diagrams. In: Rothman K, Greenland S, Lash T, editors. Modern epidemiology, 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins, 2008:183–209.
3. Lauritzen S. Causal inference from graphical models. In: Cox D, Kluppelberg C, editors. Complex stochastic systems. Boca Raton, FL: Chapman and Hall and CRC Press, 2001:63–107.

4. Pearl J. Causal diagrams for empirical research. *Biometrika* 1995;82:669–710.
5. Spirtes P, Glymour C, Scheines R. *Causation, prediction, and search*, 2nd ed. Cambridge, MA: MIT Press, 2000.
6. Greenland S, Robins J. Identifiability, exchangeability, and epidemiological confounding. *Int J Epidemiol* 1986;15:413–19.
7. Stone R. The assumptions on which causal inferences rest. *J R Stat Soc* 1993;55:455–66.
8. Pearl J. *Causality: models, reasoning, and inference*. New York: Cambridge University Press, 2000; 2nd ed., 2009.
9. Pearl J. Comment: graphical models, causality, and intervention. *Stat Sci* 1993;8:266–9.
10. Engle R, Hendry D, Richard J. Exogeneity. *Econometrica* 1983;51:277–304.
11. Rosenbaum P, Rubin D. The central role of propensity score in observational studies for causal effects. *Biometrika* 1983;70:41–55.
12. Kuroki M, Cai Z. Selection of identifiability criteria for total effects by using path diagrams. In: Chickering M, Halpern J, editors. *Uncertainty in artificial intelligence, proceedings of the twentieth conference*. Arlington, VA: AUAI, 2004:333–40.
13. Kuroki M, Miyakawa M. Covariate selection for estimating the causal effect of control plans using causal diagrams. *J Jpn R Stat Soc Ser B* 2003;65:209–22.
14. Bishop Y, Fienberg S, Holland P. *Discrete multivariate analysis: theory and practice*. Cambridge, MA: MIT Press, 1975.
15. Greenland S, Pearl J. Adjustments and their consequences – collapsibility analysis using graphical models. *Int Stat Rev* 2011;79:401–26.
16. Greenland S, Robins J, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci* 1999;14:29–46.
17. Neyman J. On the application of probability theory to agricultural experiments. *Essay on principles*. Section 9. *Stat Sci* 1923;5:465–80.
18. Rubin D. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 1974;66:688–701.
19. Pearl J. *Probabilistic reasoning in intelligent systems*. San Mateo, CA: Morgan Kaufmann, 1988.
20. Pearl J. *Causality: models, reasoning, and inference*, 2nd ed. New York: Cambridge University Press, 2009.
21. Pearl J. On a class of bias-amplifying variables that endanger effect estimates. In: Grünwald P, Spirtes P, editors. *Proceedings of the twenty-sixth conference on uncertainty in artificial intelligence*. Corvallis, OR: AUAI, 2010:417–24. Available at: [http://ftp.cs.ucla.edu/pub/stat\\_ser/r356.pdf](http://ftp.cs.ucla.edu/pub/stat_ser/r356.pdf).
22. Wooldridge J. Should instrumental variables be used as matching variables? Technical report, Michigan State University, MI, 2009. Available at: <http://econ.msu.edu/faculty/wooldridge/docs/treat1r6.pdf>.
23. Pearl J. Linear models: a useful “microscope” for causal analysis. *J Causal Inference* 2013;1:155–70.
24. Tian J, Paz A, Pearl J. Finding minimal separating sets. Technical Report R-254, Computer Science Department, University of California, Los Angeles, CA, 1998. Available: [http://ftp.cs.ucla.edu/pub/stat\\_ser/r254.pdf](http://ftp.cs.ucla.edu/pub/stat_ser/r254.pdf).
25. Bareinboim E, Pearl J. Transportability of causal effects: completeness results. In: Hoffman J, Selman B, editors. *Proceedings of the twenty-sixth AAAI conference on artificial intelligence*. Menlo Park, CA: AAAI Press, 2012:698–704.
26. Daniel RM, Kenward MG, Cousens SN, Stavola BL. Using causal diagrams to guide analysis in missing data problems. *Stat Methods Med Res* 2011;21:243–56.
27. Geneletti S, Richardson S, Best N. Adjusting for selection bias in retrospective case-control studies. *Biostatistics* 2009;10:17–31.
28. Shpitser I, VanderWeele T, Robins J. On the validity of covariate adjustment for estimating causal effects. In: Grünwald P, Spirtes P, editors. *Proceedings of the twenty-sixth conference on uncertainty in artificial intelligence*. Corvallis, OR: AUAI, 2010:527–36.
29. Robins J. Causal inference from complex longitudinal data. In: Berkane M, editor. *Latent variable modeling and applications to causality*. New York: Springer, 1997:69–117.
30. Greenland S, Pearl J, Robins J. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37–48.
31. Pearl J. Why there is no statistical test for confounding, why many think there is, and why they are almost right. Technical Report R-256, Department of Computer Science, University of California, Los Angeles, CA, 1998. Available at: [http://ftp.cs.ucla.edu/pub/stat\\_ser/R256.pdf](http://ftp.cs.ucla.edu/pub/stat_ser/R256.pdf).
32. Weng H-Y, Hsueh Y-H, Messam LL, Hertz-Picciotto I. Methods of covariate selection: directed acyclic graphs and the change-in-estimate procedure. *Am J Epidemiol* 2009;169:1182–90.
33. Evans D, Chaix B, Lobbedez T, Verger C, Flahault A. Combining directed acyclic graphs and the change-in-estimate procedure as a novel approach to adjustment-variable selection in epidemiology. *BMC Med Res Methodol* 2012;12. DOI:10.1186/1471-2288-12-156.
34. Dawid A. Conditional independence in statistical theory. *J R Stat Soc Ser B* 1979;41:1–31.
35. Lauritzen SL, Dawid AP, Larsen BN, Leimer HG. Independence properties of directed Markov fields. *Networks* 1990;20:491–505.