

# Confounding Equivalence in Observational Studies (or, when are two measurements equally valuable for effect estimation?)

Judea Pearl

Cognitive Systems Laboratory  
Computer Science Department

University of California, Los Angeles, CA 90024 USA

*judea@cs.ucla.edu*

Azaria Paz

Department of computer Science

Technion IIT, Haifa 3200

*paz@cs.technion.ac.il*

September 2008

Revised: November 2009

## Abstract

The paper provides a simple test for deciding, from a given causal diagram, whether two sets of covariates 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(s) are identical in both sets. Applications to covariate selection and model testing are discussed.

## 1 Introduction

The common method of reducing confounding bias in observational studies is to *adjust* for a set of covariates judged to be “confounders,” that is, variables capable of producing spurious associations between treatment and outcome, not attributable to their causal dependence. It is now a common understanding that 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, and that such relationships can effectively be represented in the form of directed acyclic graphs (DAG’s) (Pearl, 1995; Lauritzen, 2001; Spirtes et al., 2000; Glymour and Greenland, 2008).

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 (Stone, 1993; Greenland et al., 1999; Pearl, 2000). A graphical criterion for selecting an admissible set is given by the “back-door” test (Pearl, 1993, 2000) which was shown to entail zero-bias, or “no confoundedness,” assuming correctness of the causal assumptions encoded in the DAG. Related notions are “exchangeability” (Greenland and Robins, 1986), “exogeneity” (Engle et al., 1983), and “strong ignorability” (Rosenbaum and Rubin, 1983).

This paper addresses a different question: Given two sets of covariates 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. Second, assuming that the structure of the underlying DAG is only partially known, one may wish to assess, using  $c$ -equivalence tests, whether a 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. Finally, tests for bias equality offer computational advantages over tests for conditional independencies; characterizing graphs in term of the former, may lead to more effective ways of inferring graph structures from data.

## 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

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

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 (2000) 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  $T \approx Z$ , 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

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

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

**Definition 2.** (*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 “admissible” (for adjustment) relative to the causal effect of  $X$  on  $Y$ , if the following equality holds:

$$\sum_z P(y|x, z)P(z) = P(y|do(x)) \quad (3)$$

Equivalently, one can define admissibility using the equalities:

$$P(y|do(x)) = P(Y_x = y) = \sum_u P(y|x, u)P(u) \quad (4)$$

where  $Y_x$  is the counterfactual or “potential outcome” variable (Neyman, 1923; Rubin, 1974) and  $U$  ranges over all variables that affect either  $X$  or  $Y$ , hidden as well as observables (Stone, 1993). The equivalence of the three definitions is demonstrated in Pearl (2000).<sup>1</sup>

**Definition 3.** (*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  $Y$ ,” and then,  $X$  and  $Y$  are independent given  $S$ , written  $X \perp\!\!\!\perp Y | S$ , in every probability distribution that can be generated by a process structured along  $G$  (Pearl, 1988).

**Lemma 1.** (*The back-door criterion*)

Let  $G$  be a directed acyclic graph (DAG) that encodes the causal relationships between variables in a problem, observables as well as unobservable. A sufficient condition for a subset  $S$  of covariates to be admissible is that it satisfies the following two conditions (Pearl, 1993):

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  $X$ .

For proof and intuition behind the back-door test, especially the requirement of no descendants, see (Pearl, 2009, p. 339).

Clearly, if two subsets  $Z$  and  $T$  are admissible they must be  $c$ -equivalent, for their adjustment estimands coincide with the causal effect  $P(y|do(x))$ . Therefore, a trivial graphical condition for  $c$ -equivalence is for  $Z$  and  $T$  to satisfy the back-door criterion of Lemma 1. This condition, as we shall see in the next section, is rather weak;  $c$ -equivalence extends beyond admissible sets.

---

<sup>1</sup>The  $do(x)$  notation is preferred here because it rests on realizable variables and explicit causal processes. The expression  $\sum_u P(y|x, u)P(u)$  has the appearance of a standard statistical parameter, but is in fact a causal notion, for it assumes knowledge of “all causes” and cannot be defined in terms of distributions of observed variables (Pearl 2000, Chapter 6; 1998).

### 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 (2) as

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

and (i) further yields  $P(z|t, x) = P(z|t)$ , from which (2) follows:

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

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

$$\begin{aligned} C^* &: X \perp\!\!\!\perp Z|T \quad \text{and} \quad Y \perp\!\!\!\perp T|Z, X \\ C^{**} &: X \perp\!\!\!\perp T|Z \quad \text{and} \quad Y \perp\!\!\!\perp Z|T, X \end{aligned} \tag{5}$$

**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 admissible (by virtue of satisfying the back-door criterion); they must therefore be  $c$ -equivalent. Yet neither  $C^*$  nor  $C^{**}$  holds in this case.

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 paths in the graph.<sup>2</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

---

<sup>2</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  $W_1$  independent given  $X$  and  $W_2$ . But this is true only if the path  $(X, V_1, V_2, Y)$  is blocked.

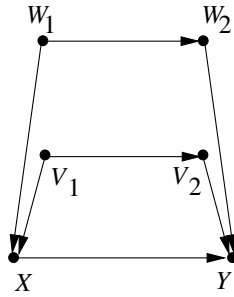


Figure 1: The sets  $T = \{V_1, W_1\}$  and  $Z = \{V_2, W_2\}$  satisfy the conditions of Theorem 1. The sets  $T = \{V_1, W_2\}$  and  $Z = \{V_2, W_2\}$  block all back-door paths between  $X$  and  $Y$ , hence they are admissible and  $c$ -equivalent. Still they do not satisfy the conditions of Theorem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

$T = \{W_1\}$  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 below) that detects *all*  $c$ -equivalent subsets in a graph.

**Definition 4.** (*Markov boundary*)

For any set of variables  $S$ , let  $S_m$  be the minimal subset of  $S$  that satisfies the condition

$$X \perp\!\!\!\perp S \mid S_m \tag{6}$$

In other words, measurement of  $S_m$  renders  $X$  independent of all other members of  $S$ , and no proper subset of  $S_m$  has this property. This minimal subset is called the *Markov Boundary* of  $X$  relative to  $S$ , or “*Markov boundary*” for short.

The Markov boundary as defined above is known to be unique for all strictly positive distributions (Pearl, 1988, p. 97). We can also define this notion relative to a graph  $G$  by interpreting the  $\perp\!\!\!\perp$  sign in Eq. (6) to denote  $d$ -separation. We will not distinguish between the graphical and statistical definitions because all our theorems would be valid in both interpretations.<sup>3</sup>

**Lemma 2.** *Every set of variables,  $S$ , is  $c$ -equivalent to its Markov boundary  $S_m$ .*

**Proof.**

Choosing  $Z = S$  and  $T = S_m$  satisfies the two conditions of Theorem 1; (i) is satisfied by the definition of  $S_m$ , while (ii) is satisfied by subsumption ( $T \subseteq Z$ ).

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

1.  $Z_m = T_m$

or

2.  $Z$  and  $T$  are admissible

---

<sup>3</sup>Note however that the graphical definition gives rise to a *unique* minimal subset  $S_m$  that satisfies Eq. (6). This follows from the intersection property of  $d$ -separation (Pearl, 1988, p. 97).

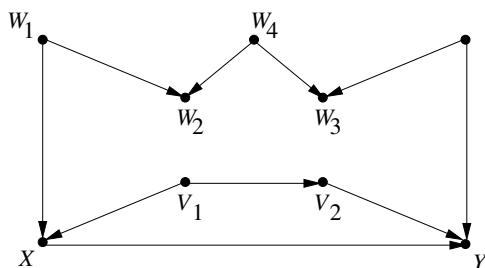


Figure 2:  $W_3$  and  $W_4$  are non-admissible yet  $c$ -equivalent; both having  $\emptyset$  as a Markov boundary. However,  $W_2$  and  $W_3$  are not  $c$ -equivalent with Markov boundaries  $W_2$  and  $\emptyset$ , respectively.

## Proof

### 1. Proof of sufficiency:

Condition 2 is sufficient since admissibility renders the two adjustment estimands in (2) equal to the causal effect. Condition 1 is sufficient by reason of Lemma 2, which yields:

$$Z \approx Z_m \approx T_m \approx T$$

### 2. Proof of necessity:

If exactly one of  $(Z, T)$  is 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 admissible or, equivalently, that none of  $Z_m$  or  $T_m$  is 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, excluding the degenerate case where  $T_m$  and  $Z_m$  are both disconnected from  $X$ , there must be a path  $p_1$  from either  $Z_m$  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, we conclude that if (1) and (2) are both violated, 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 for some values  $z'$  and  $z''$  of  $Z$  and for all  $t$  of  $T$  we have

$$\begin{aligned} P(y|z', t, x) &\neq P(y|t, x) \\ P(z''|x, t) &\neq P(z''|t) \end{aligned}$$

Thus, conditional on  $t$ ,  $Z$  acts as an instrumental variable with respect to the pair  $(X, Y)$  and, based on the examples in Wooldridge (2009), there is a parametrization of the graph that would yield the inequality  $A(x, y, T_m) \neq A(x, y, Z_m)$ . This proves the necessary part of Theorem 2.

Figure 2 illustrates the power of Theorem 2. In this model, no subset of  $\{W_1, W_2, W_3\}$  is admissible (because of the back-door path through  $V_1$  and  $V_2$ ) and, therefore, equality of Markov boundaries 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 block the same set of paths, 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 Markov boundary of each is the null set,  $\{\emptyset\}$ .

At this point it is worth noting that, based on Theorem 2,  $c$ -equivalence can be tested by a step-wise process,

$$T \approx T_1 \approx T_2 \approx \dots \approx Z$$

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.

$$T \approx \dots \approx T_m \approx \dots \approx Z_m \approx \dots \approx Z$$

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 Markov boundary  $Z_m$ ; again, in any order. Finally, the middle section, from  $T_m$  to  $Z_m$ , amounts to traversing a chain of admissible sets, using both deletion and addition of nodes, one at a time. A Theorem due to Tian et al. (1998) ensures that such a step-wise transition is always possible between any two admissible sets. In case  $T$  or  $Z$  are 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 stepwise transition from  $T = \{W_1, W_2, V_1\}$  to  $Z = \{V_2, W_3\}$ . Starting with  $T$ , we obtain:

$$T = \{W_1, W_2, V_1\} \approx \{W_2, V_1\} \approx \{V_1\} \approx \{V_1, V_2\} \approx \{V_2\} \approx \{V_2, W_3\} = Z$$

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

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

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 Markov boundaries shown in the graph.

Note also that each step in the process  $T \approx \dots \approx T_m$  (as well as  $Z_m \approx \dots \approx Z$ ) is licensed by condition (i) of Theorem 1, while each step in the intermediate process  $T_m \approx \dots \approx Z_m$  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<sup>4</sup> 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 the next section.

## 4 Discussion and refinements

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) while  $c$ -equivalence is purely statistical. Why need one resort to causal assumptions to characterize a property that relies on no

---

<sup>4</sup>A diagram  $G$  is said to be compatible with a probability function  $P$  if  $G$  can generate  $P$ , i.e., if every  $d$ -separation in  $G$  corresponds to a valid conditional independence in  $P$ .

such assumption? Evidently, the notion of admissibility as it was used in the proof of Theorem 2 was merely a surrogate carrier of statistical information; its causal reading 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 first step is to apply Theorem 1 to the special case where  $T$  is a subset of  $Z$ .

**Theorem 3.** (*Set-subset equivalence*)

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 3 generalizes similar theorems by Stone (1993) and Robins (1997), in which  $T \cup S$  is assumed to be admissible (see also (Greenland et al., 1999)). The importance of this generalization was demonstrated by several examples in Section 3. Theorem 3 on the other hand invokes only the distribution  $P(x, y, z, t)$  and makes no reference to  $P(y|do(x))$  or to the “all-causes” set  $U$ , as in Eq. (3) or (4).



Theorem 3 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 3 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  $\{T \cup S\} = \{V_1, V_2, W_1, W_2\}$

The weakness of Theorem 3 is that it is applicable to set-subset relations only. A natural attempt to generalize the theorem would be to posit the condition that  $T$  and  $Z$  each be  $c$ -equivalent to  $T \cup Z$ , and use Theorem 3 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,  $T \approx Z$  and  $T \not\approx T \cup Z$ . 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 Markov boundary (Definition 4) and Lemma 2.

**Theorem 4.** *Let  $T$  and  $Z$  be two sets of covariates, containing no descendant of  $X$  and let  $T_m$  and  $Z_m$  be their Markov boundaries. 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 3.*

### 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 3), then they are not  $c$ -equivalent to each other. We will show that using “admissibility” as an auxiliary tool. We will show that failure of  $Z_m \approx T_m \cup Z_m$  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 1), which states that, for any graph  $G$ , and any two subsets of nodes  $T$  and  $Z$ , we have:

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

Applying this to the subgraph consisting of all back-door paths from  $X$  to  $Y$ , we conclude that 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 3.

Theorem 4 reveals the statistical implications of the admissibility requirement in Theorem 2. Admissibility ensures the two  $c$ -equivalence conditions:

$$T_m \approx \{T_m \cup Z_m\} \quad (7)$$

$$Z_m \approx \{T_m \cup Z_m\} \quad (8)$$

In other words, given any DAG  $G$  compatible with the conditional independencies of  $P(x, y, t, z)$ , whenever  $Z$  and  $T$  are admissible in  $G$ , we should be able to prove Eqs. (7) and (8) using the set-subset conditions of Theorem 3. These conditions can be explicated more clearly by Theorem 4.

Equations (7) and (8) then imply

$$\{T_m \cup Z_m\}_m \perp\!\!\!\perp Y | X, Z_m \quad (9)$$

$$\{T_m \cup Z_m\}_m \perp\!\!\!\perp Y | X, T_m \quad (10)$$

We illustrate these conditions using Fig. 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 (9) and (10) are satisfied because

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

Thus, implying  $Z \approx T$ . 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

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

would not be satisfied because

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

## 5 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 2000, Chapter 6; 1998). The discussion of Section 4 shows however that the admissibility of two sets,  $T$  and  $Z$ , does have testable implications. In particular, if both sets 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. (5)). Still, Eq. (7) 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  $Z$  and  $T$  cannot both be admissible.

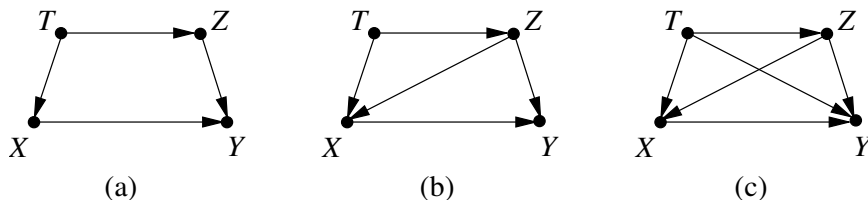


Figure 3: Failing the  $T \approx \{T \cup Z\}$  test should reject Model (a) in favor of (b) or (c). Failing  $Z \approx \{T \cup Z\}$  should reject Models (a) and (b) in favor of (c).

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

$$T_m \approx \{T_m \cup Z_m\} \approx Z_m$$

can then be used to rule out the former.

Figure 3 illustrates this possibility. Model 3(a) deems measurements  $T$  and  $Z$  as equally effective for bias removal, while models 3(b) and 3(c) deem  $T$  to be insufficient for adjustment. Submitting the data to the  $c$ -equivalence tests of Eq. (7) and Eq. (8) may reveal which of the three models should be ruled out. If both tests fail, we must rule out Models 3(a) and 3(b), while if only Eq. (8) fails, we can rule out only Model 2(a) (Eq. (7) may still be satisfied in Model 2(c) by incidental cancellation).

Of course, the same model exclusion can be deduced from conditional independence tests. For example, Models 3(a) and 3(b) both predict  $T \perp\!\!\!\perp Y | X, Z$  which, if violated in the data, would leave Model 3(c) as our choice and behoove us to adjust for both  $T$  and  $Z$ . However, such tests are generally more expensive (and less reliable) than  $c$ -equivalence test, for they require conditioning on large sets of covariates.  $c$ -equivalence tests, on the other hand, can take advantage of propensity score methods (Rosenbaum and Rubin, 1983) which reduce the dimensionality of the conditioning set to a single scalar.

This raises the interesting question of whether the discrimination power of  $c$ -equivalence equals that of conditional independence tests. We know from Theorem 4 that all  $c$ -equivalence conditions can be derived from conditional independence relations. The converse, however, is an open question.

## 6 Conclusions

Theorem 2 provides a simple graphical test for deciding whether one set of covariates has the same bias-reducing potential as another. The test requires either that both sets satisfy the back-door criterion or that their Markov boundaries be identical. Both conditions can be tested by fast, linear time algorithms, and could be used to guide researchers in deciding what measurement sets are worth performing. We have further shown that the conditions above are valid in causal as well as associational graphs; the latter can be inferred from nonexperimental data. Finally, we postulate

that  $c$ -equivalence tests could serve as valuable tools in a systematic search for graph structures that are compatible with the data.

## References

- ENGLE, R., HENDRY, D. and RICHARD, J. (1983). Exogeneity. *Econometrica* **51** 277–304.
- GLYMOUR, M. and GREENLAND, S. (2008). Causal diagrams. In *Modern Epidemiology* (K. Rothman, S. Greenland and T. Lash, eds.), 3rd ed. Lippincott Williams & Wilkins, Philadelphia, PA, 183–209.
- GREENLAND, S., PEARL, J. and ROBINS, J. (1999). Causal diagrams for epidemiologic research. *Epidemiology* **10** 37–48.
- GREENLAND, S. and ROBINS, J. (1986). Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology* **15** 413–419.
- LAURITZEN, S. (2001). Causal inference from graphical models. In *Complex Stochastic Systems* (D. Cox and C. Kluppelberg, eds.). Chapman and Hall/CRC Press, Boca Raton, FL, 63–107.
- NEYMAN, J. (1923). On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statistical Science* **5** 465–480.
- PEARL, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, San Mateo, CA.
- PEARL, J. (1993). Comment: Graphical models, causality, and intervention. *Statistical Science* **8** 266–269.
- PEARL, J. (1995). Causal diagrams for empirical research. *Biometrika* **82** 669–710.
- PEARL, J. (1998). Why there is no statistical test for confounding, why many think there is, and why they are almost right. Tech. Rep. R-256, <[http://ftp.cs.ucla.edu/pub/stat\\_ser/R256.pdf](http://ftp.cs.ucla.edu/pub/stat_ser/R256.pdf)>, Department of Computer Science, University of California, Los Angeles, Los Angeles, CA.
- PEARL, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, New York. 2nd edition, 2009.
- PEARL, J. (2009). *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge University Press, New York. Forthcoming.
- ROBINS, J. (1997). Causal inference from complex longitudinal data. In *Latent Variable Modeling and Applications to Causality* (M. Berkane, ed.). Springer-Verlag, New York, 69–117.
- ROSENBAUM, P. and RUBIN, D. (1983). The central role of propensity score in observational studies for causal effects. *Biometrika* **70** 41–55.
- RUBIN, D. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* **66** 688–701.

- SPIRITES, P., GLYMOUR, C. and SCHEINES, R. (2000). *Causation, Prediction, and Search*. 2nd ed. MIT Press, Cambridge, MA.
- STONE, R. (1993). The assumptions on which causal inferences rest. *Journal of the Royal Statistical Society* **55** 455–466.
- TIAN, J., PAZ, A. and PEARL, J. (1998). Finding minimal separating sets. Tech. Rep. R-254, [http://ftp.cs.ucla.edu/pub/stat\\_ser/r254.pdf](http://ftp.cs.ucla.edu/pub/stat_ser/r254.pdf), University of California, Los Angeles, CA.
- WOOLDRIDGE, J. (2009). Should instrumental variables be used as matching variables? Tech. Rep. <<https://www.msu.edu/~ec/faculty/wooldridge/current%20research/treat1r6.pdf>>, Michigan State University.