

# Signed directed acyclic graphs for causal inference

BY TYLER J. VANDERWEELE

*Department of Health Studies, University of Chicago*  
*5841 South Maryland Avenue, MC 2007, Chicago, IL 60637*  
vanderweele@uchicago.edu

JAMES M. ROBINS

*Departments of Epidemiology and Biostatistics, Harvard School of Public Health*  
*677 Huntington Avenue, Boston, MA 02115, USA*  
robins@hsph.harvard.edu

**Summary.** By introducing the notions of a monotonic effect, a weak monotonic effect and a signed edge, the directed acyclic graph causal framework can be extended so as to allow not only for the graphical representation of causal relations amongst variables but also for the sign of these causal relations. Results are developed relating monotonic effects to the sign of the causal effect of an intervention in the presence of intermediate variables. The incorporation of signed edges into the directed acyclic graph causal framework furthermore allows for the development of rules governing the relationship between monotonic effects and the sign of the covariance between two variables and also rules governing the sign of the bias which arises when control for confounding is inadequate.

*Keywords:* Bias; Causal inference; Confounding; Directed acyclic graphs; Structural equations.

## 1. Introduction

With very few exceptions (Lauritzen and Richardson, 2002), the use of graphical models in the field of causal inference has been restricted to directed acyclic graphs and graphs allowing for bidirected edges which represent unobserved common causes. The directed acyclic graph causal framework allows for the representation of causal and counterfactual relations amongst variables (Pearl, 1995; Robins, 1997; Pearl, 2000); the estimation of causal effects through the g-formula (Robins, 1986, 1987; Spirtes et al., 1993; Pearl, 1993); the detection of independencies through the d-separation criterion (Verma and Pearl, 1988; Geiger et al., 1990; Lauritzen et al., 1990); and the implementation of algorithms to determine whether conditioning on a particular set of variables, or none at all, is sufficient to control for confounding as well as algorithms which identify such a set of variables (Pearl 1995; Galles and Pearl, 1995; Pearl and Robins, 1995; Kuroki and Miyakawa, 1999, 2003; Geng et al., 2002; Tian and Pearl, 2002).

This paper introduces into the directed acyclic graph causal framework the notion of a monotonic effect and along with its graphical counterpart, a signed edge. Considerations of monotonicity have been incorporated into many types graphical models (Wellman, 1990; Archer and Wang, 1993; Druzdzel and Henrion, 1993; Bioch, 1998; Potharst and Feelders, 2002; van der Gaag et al., 2004). Here our focus will be the causal interpretation of monotonicity relationships within the directed acyclic graph framework. We will say that some variable has a positive (or negative) monotonic effect on another if intervening to increase the former will never for any individual decrease (or increase) the latter. By incorporating these notions of monotonic effects, the directed acyclic graph causal framework can be extended in various directions. Signs can be added to the edges of the directed acyclic graph to indicate the presence of a particular positive or negative monotonic effect. Using the signs of these edges, one may then determine the sign of the causal effect of an intervention in the presence of intermediate variables. Similarly one may also determine the sign of the covariance between various nodes on the signed causal directed acyclic graph and in certain circumstances one may determine the sign of the bias resulting when control for confounding is inadequate. These results generalize for the case of non-parametric structural equations corresponding results that are more straightforward in the multivariate normal setting. Although the utility of the results presented in this paper is in part limited by the strong assumptions which must be made about monotonic effects, we argue that a theory based only on average causal effects will not work. We furthermore provide a weaker set of conditions, which we express in terms of weak monotonic effects, under which the major theorems presented still hold.

Before formally introducing the concepts of a monotonic effect, a weak monotonic effect and a signed edge, we review definitions and results concerning causal directed acyclic graphs. Following Pearl (1995), a causal directed acyclic graph is a set of nodes  $(X_1, \dots, X_n)$  and directed edges amongst nodes such that the graph has no cycles and such that for each node  $X_i$  on the graph the corresponding variable is given by its non-parametric structural equation  $X_i = f_i(pa_i, \epsilon_i)$  where  $pa_i$  are the parents of  $X_i$  on the graph and the  $\epsilon_i$  are mutually independent. These non-parametric structural equations can be seen as a generalization of the path analysis and linear structural equation models (Pearl 1995, 2000) developed by Wright (1921) in the genetics literature and Haavelmo (1943) in the econometrics literature. Robins (1995, 2003) discusses the close relationship between these non-parametric structural equation models and fully randomized causally interpreted structured tree graphs (Robins 1986, Robins 1987). Unlike their linear

counterpart, non-parametric structural equations are entirely general –  $X_i$  may depend on any function of its parents and  $\epsilon_i$ . The non-parametric structural equations encode counterfactual relationships amongst the variables represented on the graph. The equations themselves represent one-step ahead counterfactuals with other counterfactuals given by recursive substitution. A node  $A$  will be a parent of  $Y$  if there is some level of all variables that precede  $Y$  such that intervening to set  $A$  to different levels will allow  $Y$  to vary even after intervening to fix all other variables that precede  $Y$ . If there exists some level of  $A$  such that intervening to set  $C$  to different levels will allow  $B$  to vary even after fixing  $A$  and there exists some level of  $B$  such that intervening to set  $C$  to different levels will allow  $A$  to vary even after fixing  $B$  then  $C$  is said to be a common cause of  $A$  and  $B$ . The requirement that the  $\epsilon_i$  be mutually independent is essentially a requirement that there is no variable absent from the graph which, if included on the graph, would be a parent of two or more variables (Pearl, 1995, 2000).

A path is a sequence of nodes connected by edges regardless of arrowhead direction; a directed path is a path which follows the edges in the direction indicated by the graph’s arrows. We will say that  $V_1, \dots, V_n$  constitutes an ordered list if  $i < j$  implies that  $V_i$  is not a descendent of  $V_j$ . A collider is a particular node on a path such that both the preceding and subsequent nodes on the path have directed edges going into that node i.e. both the edge to and the edge from that node have arrowheads into the node. A path between  $A$  and  $B$  is said to be blocked given some set of variables  $Z$  if either there is a variable in  $Z$  on the path that is not a collider or if there is a collider on the path such that neither the collider itself nor any of its descendants are in  $Z$ . It has been shown that if all paths between  $A$  and  $B$  are blocked given  $Z$  then  $A$  and  $B$  are conditionally independent given  $Z$  (Verma and Pearl, 1988; Geiger et al., 1990; Lauritzen et al., 1990). The directed acyclic graph causal framework has proven to be particularly useful in determining whether conditioning on a given set of variables, or none at all, is sufficient to control for confounding. The most important result in this regard is the back-door path criterion (Pearl, 1995). We denote the counterfactual value of  $Y$  intervening to set  $A = a$  by  $Y_{A=a}$ . We say that  $Z$  suffices to control for confounding for the estimation of the causal effect of  $A$  on  $Y$  if  $pr(Y_{A=a}|Z = z) = pr(Y|Z = z, A = a)$  for all  $z$  and  $a$ . A back-door path from some node  $A$  to another node  $Y$  is a path which begins with a directed edge into  $A$ . Pearl (1995) showed that for intervention variable  $A$  and outcome  $Y$ , if a set of variables  $Z$  is such that no variable in  $Z$  is a descendent of  $A$  and such that  $Z$  blocks all back-door paths from  $A$  to  $Y$  then conditioning on  $Z$  suffices to control for confounding for the estimation of the causal effect of  $A$  on  $Y$ .

We present an example taken from Greenland et al. (1999) which illustrates the use of the back-door path criterion and which also motivates the development of some of the theory in this paper. Consider a study of the relation of antihistamine treatment, denoted by  $E$ , and asthma incidence, denoted by  $D$ , among first-grade children attending various public schools. Suppose that air pollution level, denoted by  $A$ , is independent of sex, denoted by  $B$ , among first-grade public school children. Suppose further that sex influences the administration of antihistamine only through its relation to bronchial reactivity, denoted by  $C$ , but that sex directly influences asthma risk; suppose also that air pollution leads to asthma attacks only through its influence on antihistamine use and bronchial reactivity; and that there are no important confounders beyond air pollution, bronchial reactivity and sex. The causal relationships amongst these variables are then those given in Figure 1.

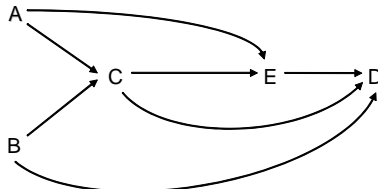


Fig. 1. Example illustrating the use of the back-door path criterion

Under the assumptions given above, conditioning on  $A$ ,  $B$  and  $C$  suffices to control for confounding; conditioning on  $A$  and  $C$  alone also suffices to control for confounding; conditioning on  $B$  and  $C$  alone also suffices to control for confounding. However conditioning on any other subset of  $A$ ,  $B$  and  $C$  does not suffice. Thus if data were available only on antihistamine use  $E$  (recorded as a

binary variable: yes or no), asthma  $D$  and sex  $B$  then we could not produce valid estimates of the causal effect of  $E$  on  $D$ . We will return to this example at the end of the paper and show that under assumptions concerning weak monotonic effects, the estimate of the causal effect of  $E$  on  $D$  controlling only for  $B$  is in fact conservative.

The remainder of the paper is organized as follows. Section 2 presents the definitions of a monotonic effect and a weak monotonic effect along with related definitions concerning the signs of a graph's edges and paths; a number of technical lemmas and a graph theoretic result are also given. Section 3 presents a result which allows for the determination of the sign of the causal effect of an intervention in the presence of intermediate variables; in this section it is also argued that a theory of signed directed acyclic graphs based only on average causal effects will fail. In section 4, a probability lemma is given which is needed to prove the section's theorem concerning the rules governing signed edges and covariance. Section 5 gives a result concerning the use of monotonic effects to determine the sign of the bias when control for confounding is inadequate. Throughout sections 3 through 5, examples are provided to illustrate the use of the sections' theorems. Some final comments are given in section 6.

## 2. Monotonic effects and signed edges

Various extensions to the directed acyclic graph causal framework are made possible by introducing the idea of a monotonic effect. The definition of a monotonic effect is given in terms of a directed acyclic graph's nonparametric structural equations.

*Definition 1 (Monotonic Effect).* The non-parametric structural equation for some node  $Y$  on a causal directed acyclic graph with parent  $A$  can be expressed as  $Y = f(pa_Y^*, A, \epsilon_Y)$  where  $pa_Y^*$  are the parents of  $Y$  other than  $A$ ; the variable  $A$  is said to have a positive monotonic effect on  $Y$  if for all  $pa_Y^*$  and  $\epsilon_Y$ ,  $f(pa_Y^*, a_1, \epsilon_Y) \geq f(pa_Y^*, a_2, \epsilon_Y)$  whenever  $a_1 \geq a_2$ ; the variable  $A$  is said to have a negative monotonic effect on  $Y$  if for all  $pa_Y^*$  and  $\epsilon_Y$ ,  $f(pa_Y^*, a_1, \epsilon_Y) \leq f(pa_Y^*, a_2, \epsilon_Y)$  whenever  $a_1 \geq a_2$ .

*Remark 1.* We first note that the definition of a monotonic effect requires that the intervention variable  $A$  and the response variable  $Y$  be ordered. The definition is inapplicable to any intervention variable or response variable that is categorical but not also ordinal (e.g. hair color, place of birth, etc.). Beyond this, the definition of a monotonic effect essentially requires that some intervention  $A$  either increase or decrease some other variable  $Y$  not merely on average over the entire population but rather for every individual in that population, regardless of the interventions made on the other parents of  $Y$ . The requirements for the attribution of a monotonic effect are thus considerable. However whenever a particular intervention is always beneficial or neutral for all individuals with respect to a particular outcome, one will be able to attribute a positive monotonic effect; whenever the intervention is always harmful or neutral for all individuals with respect to a particular outcome, one will be able to attribute a negative monotonic effect. Note the presence of a monotonic effect is relative to the parent set  $pa_Y^*$ ; thus  $A$  may have a monotonic effect on  $Y$  on a particular causal directed acyclic graph but may not on another graph which contains additional parents of  $Y$ . With data, one might be able to reject the hypothesis of a positive monotonic effect if for some  $a_1 > a_2$  the average causal effect over the population of setting  $A = a_1$  is less than that of setting  $A = a_2$ . However, because for any individual we observe the outcome only under one particular value of the intervention variable, the presence of a monotonic effect is not identifiable. We must thus rely on substantive knowledge of the problem under consideration in order to attribute a monotonic effect. The theorems presented in this paper are in fact true under slightly weaker conditions which are identifiable when data on all of the directed acyclic graph's variables are observed. We thus introduce the concept of a weak monotonic effect.

*Definition 2 (Weak Monotonic Effect).* A parent  $A$  of some node  $Y$  on a causal directed acyclic graph is said to have a weak positive monotonic effect on  $Y$  if the survivor function  $S(y|pa_Y^*, a) = pr(Y \geq y|pa_Y^*, a)$  is such that whenever  $a_1 \geq a_2$  we have  $S(y|pa_Y^*, a_1) \geq S(y|pa_Y^*, a_2)$  for all  $y$  and all  $pa_Y^*$ ; the variable  $A$  is said to have a weak negative monotonic effect on  $Y$  if  $S(y|pa_Y^*, a_1) \leq S(y|pa_Y^*, a_2)$  whenever  $a_1 \geq a_2$  for all  $y$  and all  $pa_Y^*$ .

*Remark 2.* Note that the definition of a weak monotonic effect can be stated simply as a condition about stochastic dominance. The cumulative distribution function  $F_X$  is said to stochastically dominate  $F'_X$  to the first order if for every  $x$  the probability that  $X \leq x$  under  $F_X$  is less than or equal to the probability that  $X \leq x$  under  $F'_X$ . Thus  $A$  has a weak positive monotonic effect on  $Y$  if for all  $pa_Y^*$  the distribution of  $Y$  conditional on  $pa_Y^*$  and  $A = a_1$  stochastically dominates to the first order the distribution of  $Y$  conditional on  $pa_Y^*$  and  $A = a_2$  whenever  $a_1 \geq a_2$ . Note further that the definition of a weak monotonic effect is given in terms of conditional probabilities and makes no reference to counterfactuals (cf. Dawid 2000, 2002). Because the conditional probabilities are identifiable from data, the presence of a weak monotonic effect is also identifiable. A weak monotonic effect is not only identifiable but it also constitutes a substantially less stringent condition. In the case of a binary outcome  $Y$  all that is required for a weak monotonic effect is that a higher value of  $A$  makes the outcome  $Y$  at least as likely regardless of the value of the parents of  $Y$  other than  $A$ . When  $Y$  is not binary the presence of a weak monotonic effect is equivalent to the statement that for all  $y$  a higher value of  $A$  makes the event  $\{Y \geq y\}$  as likely or more likely regardless of the value of the parents of  $Y$  other than  $A$ . If intervening to increase  $A$  led to a decrease in  $Y$  for only a single individual the strong conditions for a monotonic effect would fail. The less stringent conditions required for attributing a weak monotonic effect circumvent this difficulty. Consider, for example, an analysis comparing the effect on thyroid cancer of no radiation exposure to a high level of radiation exposure. For most individuals the exposure to a high level of radiation will increase the likelihood of developing thyroid cancer. However, exposure to a high level of radiation may, for a few individuals, destroy already existing thyroid cancer cells and thereby prevent the cancer's development. Within joint strata of particular sets of background variables on a causal directed acyclic graph, the exposure to radiation will increase the overall likelihood of thyroid cancer but it may not do so for every individual in the population. In such a scenario the high level of radiation exposure would not have a monotonic effect on the development of thyroid cancer but it would have a weak monotonic effect. Note the presence of a weak monotonic effect is also relative to the parent set  $pa_Y^*$ . It was noted above that the presence of a weak monotonic effect is identifiable. Although it is thus possible to empirically verify the presence of a weak monotonic effect, in practice a researcher will still likely rely on substantive knowledge of the problem in attributing a weak monotonic effect. Unless the variable  $Y$  and all its parents are ordinal with a small number of levels, a very large data set would be required to empirically verify the presence of a weak monotonic effect. Furthermore, the presence of a weak monotonic effect is only identifiable when data are observed on  $Y$  and all of its parents; however, as will be seen below, the theorems given in the paper are principally useful precisely when some of this data is missing.

*Remark 3.* The idea of a weak positive monotonic effect is closely associated with that of positive qualitative influence in Wellman's qualitative probabilistic networks (Wellman, 1990). A weak positive monotonic effect and positive qualitative influence coincide for parent  $A$  and child  $Y$  when the "context" for qualitative influence is chosen to be the parents of  $Y$  other than  $A$ . Both ideas concern stochastic dominance. If hypothetical interventions are conceivable only for certain nodes on a graph then it may be more appropriate to speak of "qualitative positive influence" than of "weak monotonic effects" for those variables for which no intervention is possible. We provide a few additional comments relating our work to that of Wellman (1990). A simplified version of Lemma 4 below and the first half of the statement of Theorem 2 follow in a fairly straightforward manner from Wellman's Theorems 4.2 and 4.3. Lemma 4 and Theorem 2, however, give the results in much greater generality. Theorem 1 and Theorems 3-5 in the present paper all contain wholly new material. Wellman (1990) gives considerable attention to the preservation of monotonicity under edge reversal, to the necessity of first order stochastic dominance for "propagating influences" and to the "propagation" of sub-additive and super-additive relationships on probabilistic networks.

A monotonic effect is a relation between two nodes on a directed acyclic graph and as such it is associated with an edge. The definition of the sign of an edge can be given either in terms of monotonic effects or weak monotonic effects. We can define the sign of an edge as the sign of the monotonic effect or weak monotonic effect to which the edge corresponds; this in turn gives rise

to a natural definition for the sign of a path.

*Definition 3 (Sign of an Edge).* An edge from  $A$  to  $Y$  on a causal directed acyclic graph is said to be of positive sign if  $A$  has a (weak) positive monotonic effect on  $Y$ . An edge from  $A$  to  $Y$  is said to be of negative sign if  $A$  has a (weak) negative monotonic effect on  $Y$ . If  $A$  has neither a (weak) positive monotonic effect nor a (weak) negative monotonic effect on  $Y$ , then the edge from  $A$  to  $Y$  is said to be without a sign.

*Definition 4 (Sign of a Path).* The sign of a path on a causal directed acyclic graph is the product of the signs of the edges that constitute that path. If one of the edges on a path is without a sign then the sign of the path is said to be undefined.

We will call a causal directed acyclic graph with signs on those edges which allow them a signed causal directed acyclic graph. The theorems in this paper are given in terms of signed paths so as to be applicable to both monotonic effects and weak monotonic effects. Theorems 3, 4 and 5 are proved for the case of weak monotonic effects and the case of monotonic effects thereby follows immediately since the presence of a monotonic effect clearly implies the presence of a weak monotonic effect. Theorem 2, however, contains statements about monotonic effects in the conclusion of the Theorem and not just in the antecedent; Theorem 2 thus must be proved for the case of monotonic effects and weak monotonic effects separately. Theorem 1 is a purely graph-theoretic result and as such does not concern monotonic effects. The statements about signed paths may thus be interpreted throughout as corresponding to either monotonic effects or weak monotonic effects. One further definition will be useful in Theorem 4 concerning the sign of the covariance of the graph's variables.

*Definition 5 (Monotonic Association).* Two variables  $A$  and  $Y$  are said to be positively monotonically associated if all directed paths between  $A$  and  $Y$  are of positive sign and all common causes  $C_i$  of  $A$  and  $Y$  are such that all directed paths from  $C_i$  to  $A$  are of the same sign as all directed paths from  $C_i$  to  $Y$ ; the variables  $A$  and  $Y$  are said to be negatively monotonically associated if all directed paths between  $A$  and  $Y$  are of negative sign and all common causes  $C_i$  of  $A$  and  $Y$  are such that all directed paths from  $C_i$  to  $A$  are of the opposite sign as all directed paths from  $C_i$  to  $Y$ .

Before proceeding to the central results governing monotonic effects, causal effects, covariance and confounding, we present two lemmas which will be useful throughout the development of the theory. The proofs follow almost immediately from the definitions and are suppressed.

*Lemma 1.* If  $A$  has a positive monotonic effect on  $Y$  then  $-A$  has a negative monotonic effect on  $Y$ ; and  $A$  has a negative monotonic effect on  $-Y$ . If  $A$  has a negative monotonic effect on  $Y$  then  $-A$  has a positive monotonic effect on  $Y$ ; and  $A$  has a positive monotonic effect on  $-Y$ .

*Lemma 2.* If on a path between  $A$  and  $B$  there is a node  $V$  other than  $A$  or  $B$  then replacing  $V$  by its negation  $-V$  does not change the sign of that path.

The application of these two lemmas can be further extended by the graph theoretic result given in Theorem 1. The proof of Theorem 1 and those of all subsequent results are given in the Appendix unless otherwise indicated. Full details of proofs given in outline are available from the authors upon request.

*Theorem 1.* If the sign of every directed path from  $A$  to  $B$  is positive then there exist nodes  $W_1, \dots, W_t$  on directed paths between  $A$  and  $B$  such that if  $W_1, \dots, W_t$  are replaced by their negations then the sign of every edge on every directed path from  $A$  to  $B$  is positive.

In subsequent sections we will also need a number of technical lemmas in order to prove the paper's major theorems. A weak monotonic effect is defined in terms of survivor functions and it is thus useful to develop certain results concerning the survivor functions of variables governed by

a causal directed acyclic graph. We will assume throughout the remainder of this paper that the variables under consideration satisfy regularity conditions that allow for the integration by parts required in the proof of Lemma 3 (which is in turn used in the proofs of most subsequent results). If conditional cumulative distribution functions are continuously differentiable then the regularity conditions will be satisfied. Härdle et al. (1998, p72) give relatively weak conditions under which such integration by parts is possible; alternatively, the existence of the Lebesgue-Stieltjes integrals found in the proof of Lemma 3 also suffices to allow integration by parts. Note that Lemma 3 will always be applied either to the function  $h(x) = x$  or to conditional survivor functions so the requirement that the variables under consideration satisfy regularity conditions that allow for the integration by parts required in the proof of Lemma 3 is simply a restriction on the distributions of the random variables. Lemma 3 relates non-decreasing functions of random variables and non-decreasing survivor functions to non-decreasing conditional expectations. Lemma 4 relates weak monotonic effects to non-decreasing survivor functions and non-decreasing conditional expectations and is used in the proofs of the theorems throughout this paper. A generalization of Lemma 4, which is needed in the proofs of Theorems 4 and 5, is stated and proved in the Appendix.

*Lemma 3.* If  $h(y, a, r)$  is non-decreasing in  $y$  and in  $a$  and  $S(y|a, r) = pr(Y > y|A = a, R = r)$  is non-decreasing in  $a$  for all  $y$  then  $E\{h(Y, A, R)|A = a, R = r\}$  is non-decreasing in  $a$ .

*Corollary.* Suppose that the  $A \rightarrow Y$  edge, if it exists, is positive. Let  $X$  denote some set of non-descendants of  $Y$  that includes  $pa_Y^*$ , the parents of  $Y$  other than  $A$ , then  $E(Y|X = x, A = a)$  is non-decreasing in  $a$  for all values of  $x$ .

*Lemma 4.* Let  $X$  denote some set of non-descendants of  $A$  that blocks all backdoor paths from  $A$  to  $Y$ . If all directed paths from  $A$  to  $Y$  are positive then  $S(y|a, x)$  and  $E(y|a, x)$  are non-decreasing in  $a$ .

### 3. Monotonic effects and the sign of the causal effect of an intervention

Having established the properties of the previous section, we can now prove a result related to the preservation of monotonic effects and weak monotonic effects when marginalizing over certain variables on a causal directed acyclic graph. This theorem allows us in turn to prove a result relating monotonic effects to the sign of the causal effect of an intervention in the presence of intermediate variables. Theorem 2 is stated in terms of positive directed paths but has an obvious analogue for negative directed paths. The theorem will be used in the proofs of subsequent theorems. We note that the first part of the theorem follows immediately from Lemma 4; it also follows immediately from repeated application of Theorems 4.2 and 4.3 in Wellman (1990) or more directly from the work of Druzdzel and Henrion (1993, Theorem 4) who build on Wellman's results. The second part of the theorem is a generalization of a monotonicity result given by Cox and Wermuth (2003).

*Theorem 2.* If the sign of every directed path from  $A$  to  $Y$  on a causal directed acyclic graph  $G$  is positive then the  $A \rightarrow Y$  edge is positive on the causal directed acyclic graph  $H$  formed by marginalizing  $G$  over all variables that are not ancestors of either  $A$  or  $Y$  and all variables on directed paths between  $A$  and  $Y$ ; furthermore the  $A \rightarrow Y$  edge is positive on the causal directed acyclic graph  $J$  formed by marginalizing  $H$  over the ancestors of  $A$  or of  $Y$  which are not common causes of  $A$  and  $Y$ .

*Remark 4.* In the simple case of only a single path between two variables, say  $A$  and  $C$ , Theorem 2 implies that the relation of manifesting a positive monotonic effect or a weak positive monotonic effect is transitive. For example, if there is no direct  $A \rightarrow C$  edge, and if  $A$  has a positive monotonic effect on  $B$  and  $B$  has a positive monotonic effect on  $C$  then it follows that  $A$  has a positive monotonic effect on  $C$  on the causal directed acyclic graph with only  $A$  and  $C$ . We will show below that if the definition of a monotonic effect is made to depend on only average causal effects, this transitivity property fails to hold. One further comment about the transitivity of monotonic effects merits attention. The definition of a positive monotonic effect required  $f(pa_Y^*, a_1, \epsilon_Y) \geq$

$f(pa_Y^*, a_2, \epsilon_Y)$  whenever  $a_1 \geq a_2$ ; it did not require that the inequality hold strictly for any  $a_1, a_2$ . Of course if for all values of  $pa_Y^*, \epsilon_Y$  we had  $f(pa_Y^*, a_1, \epsilon_Y) = f(pa_Y^*, a_2, \epsilon_Y)$  for all  $a_1, a_2$  then we generally would not attribute any causal effect of  $A$  on  $Y$ ; in such circumstances, on the causal directed acyclic graph,  $A$  would not be a parent of  $Y$ . However, if the attribution of a monotonic effect required that the inequality hold strictly for some  $a_1, a_2$  then the presence of a monotonic effect would not be a transitive relation. It is possible to construct examples in which  $A$  has a positive monotonic effect on  $B$  with the inequality holding strictly for at least two values of  $A$  and in which  $B$  has a positive monotonic effect on  $C$  with the inequality holding strictly for at least two values of  $B$  but for which  $A$  has no causal effect on  $C$  (i.e. intervening to set  $A$  to any value would always leave  $C$  unchanged). However, such examples in which causal effects are not transitive are a feature of causal directed acyclic graphs generally and are not unique to the setting of monotonic effects.

We can now state and prove the result relating monotonic effects and weak monotonic effects to the sign of the causal effect of an intervention in the presence of intermediate variables.

*Theorem 3.* If  $A$  is an ancestor of  $Y$  and the sign of every directed path from  $A$  to  $Y$  is positive then  $E(Y_{A=a})$  is non-decreasing in  $a$ .

Theorem 3 states that if the sign of every directed path from  $A$  to  $Y$  is positive then intervening to increase  $A$  will always increase or leave unchanged the average value of  $Y$  over the population. The theorem also has an obvious analogue if the sign of every directed path from  $A$  to  $Y$  is negative rather than positive. Note that the theorem requires that the sign of every directed path from  $A$  to  $Y$  is positive (or negative). If two directed paths from  $A$  to  $Y$  are of different sign or if any edge on some directed path from  $A$  to  $Y$  is without sign, we cannot determine the sign of the causal effect of an intervention from Theorem 3.

*Example 1.* We illustrate the use of the theorem by considering the signed causal directed acyclic graph given in Fig. 2. Note that no sign is present on the edge  $A \rightarrow E$ .

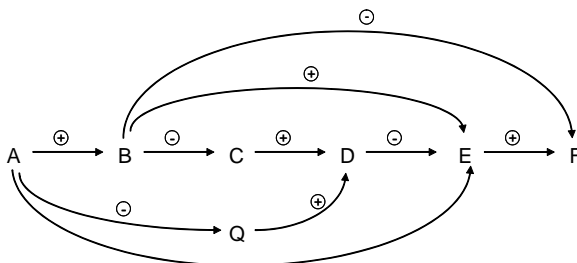


Fig. 2. Example illustrating the relationship between monotonic effects and the sign of the causal effect of an intervention

By Theorem 3, intervening to increase  $B$  will increase the average value of  $E$  over the population since all directed paths from  $B$  to  $E$  (i.e.  $B \rightarrow E$  and  $B \rightarrow C \rightarrow D \rightarrow E$ ) are of positive sign. Intervening to increase  $A$  will decrease the average value of  $D$  over the population since all directed paths from  $A$  to  $D$  (i.e.  $A \rightarrow Q \rightarrow D$  and  $A \rightarrow B \rightarrow C \rightarrow D$ ) are of negative sign. However we cannot determine from the signed causal directed acyclic graph whether intervening to increase  $A$  will increase or decrease the average value of  $E$  over the population because the  $A \rightarrow E$  edge is without a sign. Also we cannot determine from the signed causal directed acyclic graph whether intervening to increase  $B$  will increase or decrease the average value of  $F$  over the population because the paths  $B \rightarrow C \rightarrow D \rightarrow E \rightarrow F$  and  $B \rightarrow E \rightarrow F$  are of positive sign but the path  $B \rightarrow F$  is of negative sign.

Signs have sometimes informally been given to edges on a causal directed acyclic graph when intervening on the parent increases the average value of the child over the population. However when signs are given to edges in this informal manner, there are cases in which the sign from  $A$  to  $B$  might be positive and the sign from  $B$  to  $C$  might be positive (i.e. when all directed paths from

$A$  to  $C$  informally have positive sign) but intervening to increase  $A$  in fact decreases  $C$  on average over the population. In fact, even very slight departures from the requirements of a monotonic effect suffice to give counterintuitive examples. Even when intervening to increase  $A$  will increase  $B$  for all individuals in a population with the exception of only a single individual and intervening to increase  $B$  will increase or leave unchanged  $C$  for every individual in the population we may still have cases in which intervening to increase  $A$  will in fact decrease  $C$  on average over the population. Example 3 below illustrates such a case. The example illustrates first the danger of informally giving signs to the edges on a causal directed acyclic graph when intervening on the parent increases the average value of the child over the population. The example furthermore demonstrates that even very slight departures from the requirements of a monotonic effect render false the conclusions of Theorems 2 and 3 (and consequently also Theorems 4 and 5 below). We note that in the case of linear structural equations and multivariate normality, the requirements for a weak monotonic effect are satisfied and so counterintuitive situations, like that illustrated in Example 2, do not arise. Such counterintuitive also do not arise when all variables are binary because, as noted in Remark 2 above, for binary variables a non-decreasing average causal effect (conditional on the other parents, if any) implies a weak monotonic effect.

*Example 2.* Consider the causal directed acyclic graph given in Figure 3 with positive signs given informally to the  $A \rightarrow B$  and the  $B \rightarrow C$  edges. We denote these informal signs by a plus sign in quotation marks.

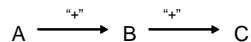


Fig. 3. Example illustrating the informal use of positive and negative signs on edges

Suppose that the variables  $A$  and  $C$  are binary and  $B$  takes on values in the set  $\{0, 1, 2\}$ . Suppose further that  $pr(A = 1) = pr(A = 0) = 1/2$ , that the structural equations for  $B$  and  $C$  are as follows:

$$\begin{aligned} B &= A + (1 - A) * 2\epsilon_B \\ C &= 1(B = 2) \end{aligned}$$

where  $\epsilon_B$  is Bernoulli with probability  $p < 1/2$  and  $1(B = 2)$  is the indicator function for  $B = 2$ . It follows that  $E[B_{A=1}] = 1$  and  $E[B_{A=0}] = 2p < 1$ . Furthermore,  $E[C_{B=0}] = E[C_{B=1}] = 0$  and  $E[C_{B=2}] = 1$ . Finally, we also have that  $E[C_{A=0}] = p$  and  $E[C_{A=1}] = 0$ . We thus have that increasing  $A$  increases  $B$  on average and increasing  $B$  increases  $C$  on average but in this example, intervening to increase  $A$  decreases  $C$  on average. Suppose that in a population of  $n$  individuals,  $p = 1/n$  then even though the departure from the requirements of a monotonic effect concerns only a single individual, intervening to increase  $A$  from 0 to 1 will still decrease  $C$  on average over the entire population.

We note that the lack of transitivity when signs are used to indicate average causal effects did not, in the example above, depend on the exact values of the distributions under consideration i.e. it arises for any  $p < 1/2$ . In fact, Wellman (1990) has shown that first order stochastic dominance (i.e. his qualitative influence) is the weakest possible relation between cumulative distribution functions for the transitivity of signs to hold.

#### 4. Covariance and monotonic effects

The notions of monotonic effects and weak monotonic effects introduced above can be used to develop rules that govern monotonic effects and covariance. When the signs of directed paths relating two variables satisfy certain conditions, it is possible to determine the sign of the covariance of these variables. However, in order to prove the central result which provides these rules we need to make use of an additional probability lemma presented below. Lemma 5 is essentially a restatement of Theorem 2.1 in Esary et al. (1967).

*Lemma 5.* Let  $f$  and  $g$  be functions with  $n$  real-valued arguments such that both  $f$  and  $g$  are non-decreasing in each of their arguments. If  $X = (X_1, \dots, X_n)$  is a multivariate random

variable with  $n$  components such that each component is independent of the other components then  $\text{cov}\{f(X), g(X)\} \geq 0$ .

This probability lemma above allows us to prove Theorem 4 concerning the rules governing covariance and monotonic effects. Theorem 4 may be seen as a generalization of the corresponding result for recursive linear structural equation models which would follow quite simply from elementary path analysis (Duncan 1975).

*Theorem 4.* If  $A$  and  $Y$  are positively monotonically associated then  $\text{Cov}(A, Y) \geq 0$ . If  $A$  and  $Y$  are negatively monotonically associated then  $\text{Cov}(A, Y) \leq 0$ .

Theorem 4 concerns the sign of covariances. In related work (VanderWeele and Robins, 2007) we have derived rules governing the sign of conditional covariance in the presence of monotonic effects; however these rules require the development of theory concerning minimal sufficient causation and assumptions beyond simply that of the presence of monotonic effects. If the directed paths relating two variables satisfy the conditions of Theorem 4, then this yields implications concerning the sign of the covariance of these variables. Obviously these conditions will not always hold; it will not always be possible to determine the sign of the covariance between any two variables simply from the signed causal directed acyclic graph. For a particular signed causal directed acyclic graph, Theorem 4 will yield implications concerning the signs of covariances for some pairs of variables and may fail to do so for others. Theorem 4 may fail to yield implications concerning the signs of covariances either because the signs of different directed paths are not congruent in the way required by the conditions of the theorem or because certain edges are without sign.

*Example 3.* Consider the signed directed acyclic graph given in Fig. 4 and note that no sign is present on the edge  $A \rightarrow F$ .

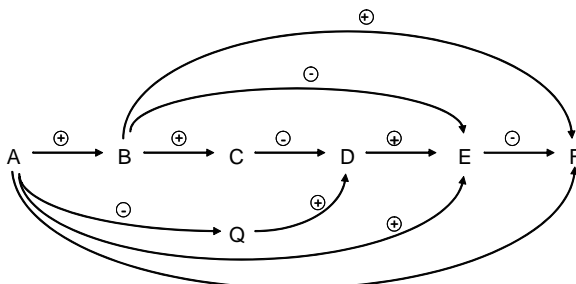


Fig. 4. Example illustrating the relationship between covariance and monotonic effects

By Theorem 4, the covariance between  $B$  and  $C$  will be positive since the sign of the only directed path from  $B$  to  $C$  is positive and  $B$  and  $C$  have no common causes except through  $B$ . The covariance between  $C$  and  $D$  will be negative since the sign of the only directed path from  $C$  to  $D$  is negative and because for the only common cause of  $C$  and  $D$ , namely  $A$ , all directed paths from  $A$  to  $C$  are positive and thus of the opposite sign as all directed paths from  $A$  to  $D$  which are negative. The covariance between  $E$  and  $F$  is undetermined because  $A$  is a common cause of  $E$  and  $F$  and the path from  $A$  to  $F$  consisting of the edge  $A \rightarrow F$  is without sign. The covariance between  $D$  and  $E$  is also undetermined because the  $D \rightarrow E$  edge is positive but  $A$  is a common cause of  $D$  and  $E$  and sign of the directed path  $A \rightarrow B \rightarrow C \rightarrow D$  from  $A$  to  $D$  is negative and so of opposite sign of the directed path from  $A$  to  $E$  consisting of the edge  $A \rightarrow E$  which is positive.

## 5. Confounding and monotonic effects

Our final result relates monotonic effects with the sign of the bias which results when control for confounding is inadequate. Manski (1997) has considered bounds for the treatment effect when the outcome is a monotonic function of the treatment. Here we consider monotonic relationships between confounding variables and the treatment and outcome. Pearl (1995) showed that for intervention variable  $A$  and outcome  $Y$ , if a set of variables  $Z$  is such that no variable in  $Z$  is a

descendent of  $A$  and  $Z$  blocks all back-door paths from  $A$  to  $Y$  then the causal effect on  $Y$  of intervening to set  $A = a$  can be determined by

$$E(Y_{A=a}) = \sum_z E(Y|A = a, Z = z)pr(Z = z).$$

Note that this result is a graphical generalization of Theorem 4 of Rosenbaum and Rubin (1983a) and of the g-formula (Robins 1986, 1987; Spirtes et al. 1993; Pearl 1993). Now if  $X$  is some set of variables that does not block all backdoor paths from  $A$  to  $Y$  and an attempt is made by means of Pearl’s result to estimate the causal effect on  $Y$  of intervening to set  $A = a$  controlling only for  $X$ , one would obtain

$$\sum_x E(Y|A = a, X = x)pr(X = x).$$

This latter expression will in general differ from the true causal effect. Theorem 5 relates monotonic effects and weak monotonic effects to the sign of the bias which arises when control for confounding is inadequate. The theorem requires that the intervention variable  $A$  be binary. We provide further discussion of this requirement following the statement of the theorem.

*Theorem 5.* Suppose that for some binary intervention  $A$  and some outcome  $Y$ , some set  $X$  of non-descendants of  $A$  does not block all backdoor paths from  $A$  to  $Y$  but does not open any backdoor paths from  $A$  to  $Y$  which were blocked without conditioning on  $X$ . Let  $S_a = \sum_x E(Y|A = a, X = x)pr(X = x)$ . If all unblocked backdoor paths from  $A$  to  $Y$  are of positive sign then  $S_1 \geq E(Y_{A=1})$  and  $S_0 \leq E(Y_{A=0})$ . If all unblocked backdoor paths from  $A$  to  $Y$  are of negative sign then  $S_1 \leq E(Y_{A=1})$  and  $S_0 \geq E(Y_{A=0})$ .

Note that if the conditions of Theorem 5 are satisfied then all unblocked backdoor paths conditioning on  $X$  are of the form  $A \leftarrow \dots \leftarrow U \rightarrow \dots \rightarrow Y$  for some  $U$  not in  $X$ . It follows from Theorem 5 that if all such unblocked backdoor paths from  $A$  to  $Y$  are of positive sign then we will overestimate the true causal risk difference and if all unblocked backdoor paths from  $A$  to  $Y$  are of negative sign then we will underestimate the true causal risk difference. Control for confounding is often inadequate when certain variables which are known to be confounders are not measured in a particular study. Various attempts have been made to provide sensitivity bounds for unmeasured confounders (Rosenbaum and Rubin, 1983b; Lin et al., 1998; Brumback et al., 2004). Theorem 5 allows the researcher in certain circumstances to determine the sign of the bias, thereby making clear whether the estimate under consideration is generous or conservative due to lack of control for certain confounding variables.

*Example 4.* We return to the example given in the Introduction illustrated by Figure 1 above. Recall that antihistamine treatment is denoted by  $E$ , asthma incidence by  $D$ , air pollution levels by  $A$ , sex by  $B$  and bronchial reactivity by  $C$ . Data is available only on antihistamine use  $E$  (recorded as a binary variable: yes or no), asthma  $D$  and sex  $B$ . Conditioning only on  $B$  does not suffice to produce valid estimates of the causal effect of  $E$  on  $D$ . In this example it is arguably quite reasonable to assume that air pollution has a weak positive monotonic effect on bronchial reactivity (relative to gender i.e. the other parent of  $C$ ) and on antihistamine use (relative to bronchial reactivity) and that bronchial reactivity has a weak positive monotonic effect on antihistamine use (relative to air pollution) and on asthma (relative to antihistamine use and gender). We may then add to Figure 1 the positive signs indicated in Figure 5.

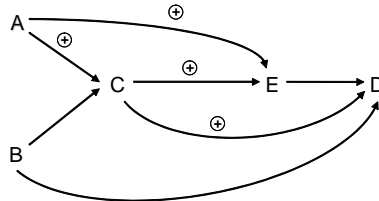


Fig. 5. Example illustrating the relationship between bias, confounding and monotonic effects

Conditioning on only  $B$ , there are two unblocked backdoor paths from  $E$  to  $D$ :  $E \leftarrow C \rightarrow D$  and  $E \leftarrow A \rightarrow C \rightarrow D$ . Both of these unblocked backdoor paths from  $E$  to  $D$  are positive. Thus

we could conclude from Theorem 5 that the estimate of the risk difference controlling only for  $B$  is an overestimate of the true causal risk difference. If the estimate controlling only for  $B$  were negative we could then conclude that the estimate was conservative and thus that antihistamine use did indeed lower asthma risk. We use simulated data to further illustrate the application of Theorem 5. In the example we will assume that  $D$  is recorded as a continuous variable (asthma severity score) and use linear regression in order to make the application of Theorem 5 particularly simple. Theorem 5 is equally applicable to binary outcomes under logistic regression; however the computation of the true expected counterfactual outcomes  $E(Y_{A=a}) = \sum_z E(Y|A = a, Z = z)pr(Z = z)$  is less straightforward due to the logit transformation. Suppose that data is generated by:

$$\begin{aligned} A &= \epsilon_A \text{ where } \epsilon_A \sim U(0.5, 1) \\ B &= \epsilon_B \text{ where } \epsilon_B \sim Ber(0.5) \\ C &= 0.5 + 0.3A + 0.02B + \epsilon_C \text{ where } \epsilon_C \sim N(0, 0.05) \\ E &= 1(0.3A + 0.4C > \epsilon_E) \text{ where } \epsilon_E \sim U(0, 1) \\ D &= 60C + 5B - 8E + \epsilon_D \text{ where } \epsilon_D \sim N(0, 10) \end{aligned}$$

and that data is available for 2000 subjects but only for the variables  $B$ ,  $E$  and  $D$ . Note that the structural equations are such that  $A$  has a positive monotonic effect on  $C$  and  $E$  and that  $C$  has a positive monotonic effect on  $E$  and  $D$ . Clearly the true causal effect,  $E(D_{E=1}) - E(D_{E=0})$ , in this example is  $-8$ . Data simulated from the structural equations above gives the following: a regression of  $D$  on  $B$  and  $E$  gives a coefficient estimate for  $E$  of  $-6.61$  with a 95% confidence interval of  $(-7.75, -5.67)$  and this estimate coincides almost exactly with the estimate of  $S_1 - S_0 = \sum_b \{E(Y|B = b, E = 1) - E(Y|B = b, E = 0)\}pr(B = b)$  calculated by exact stratification. Thus, in this example, we conclude that antihistamine use lowered asthma from data just on  $B$ ,  $E$  and  $D$  since it is a priori known that the estimate controlling only for  $B$  is conservative. If in this example data were available for  $A$  and  $C$  as well, a regression of  $D$  on  $A$ ,  $B$ ,  $C$  and  $E$  with this simulated data gives a coefficient estimate for  $E$  of  $-7.76$  with 95% confidence interval of  $(-8.65, -6.87)$ .

Note that Theorem 5 requires that the exposure variable under consideration be binary. The theorem also holds when  $A = 1$  is replaced with the maximum value of  $A$  and if  $A = 0$  is replaced with the minimum value of  $A$ . However counterexamples can be constructed to demonstrate that the theorem cannot be generalized beyond the extreme values of the intervention variable.

*Example 5.* Suppose that  $A$  were ternary and took on values in the set  $\{1, 2, 3\}$  and suppose that a single binary variable  $C$  confounded the relationship between  $A$  and  $Y$  so that the causal relationships were given by the causal directed acyclic graph given in Fig. 6.

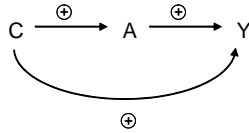


Fig. 6. Example demonstrating that Theorem 5 may fail when  $A$  is not binary

If  $f_Y(A = 3, c) = 1$  for  $c = 0, 1$  and  $f_Y(A = 1, c) = 0$  for  $c = 0, 1$  and  $f_Y(A = 2, C = 1) = 1$  and  $f_Y(A = 2, C = 0) = 0$  then both  $A$  and  $C$  will have a positive monotonic effect on  $Y$ . Suppose  $pr(C = 0) = pr(C = 1) = 1/2$ ; and  $pr(\epsilon_A = 1) = pr(\epsilon_A = 2) = pr(\epsilon_A = 3) = 0.01$  and  $pr(\epsilon_A = 4) = 0.97$ ; and suppose also  $f_A(c, \epsilon_A = 1) = 3$  for  $c = 0, 1$ ,  $f_A(c, \epsilon_A = 2) = 1$  for  $c = 0, 1$ ,  $f_A(C = 1, \epsilon_A = 3) = 3$ ,  $f_A(C = 0, \epsilon_A = 3) = 2$ ,  $f_A(C = 1, \epsilon_A = 4) = 2$  and  $f_A(C = 0, \epsilon_A = 4) = 1$  then  $C$  also has a positive monotonic effect on  $A$ . We then have  $pr(C = 1|A = 2) = pr(C = 1, A = 2)/pr(A = 2) = (0.97)(0.5)/\{(0.97)(0.5) + (0.01)(0.5)\} = 97/98$  and so  $pr(C = 0|A = 2) = 1/98$ . Thus the true causal effect on  $Y$  of intervening to set  $A = 2$  will be given by  $\sum_c E(Y|A = 2, C = c)pr(C = c) = (1)(0.5) + (0)(0.5) = 0.5$  whereas the estimate of

this causal effect without controlling for  $C$  will be given by  $E(Y|A = 2) = \sum_c E(Y|A = 2, C = c)pr(C = c|A = 2) = (1)(97/98) + (0)(1/98) = 97/98$ . In this case the estimate of the causal effect without controlling for  $C$  exceeds the true causal effect on  $Y$  of intervening to set  $A = 2$ . If however the probabilities for  $\epsilon_A$  had been  $pr(\epsilon_A = 1) = pr(\epsilon_A = 2) = pr(\epsilon_A = 4) = 0.01$  and  $pr(\epsilon_A = 3) = 0.97$  then we would still have the positive monotonic effects indicated in Fig. 6 but the estimate of the causal effect on  $Y$  of intervening to set  $A = 2$  without controlling for  $C$  would then be  $1/98$  which would be less than the true causal effect of  $0.5$ . For intermediate values of the intervention variable we thus see that the bias when control for confounding is inadequate may be in either direction even in the presence of monotonic effects.

Although Theorem 5 does not hold for intermediate values of the intervention variable, the theorem can still be useful when the intervention variable  $A$  is ordinal or continuous. A non-binary intervention variable may be dichotomized at various cut-off points. The analysis may proceed with this dichotomized intervention variable with Theorem 5 employed to assess the sign of the bias. The analysis may then be repeated a number of times at different dichotomization points and conclusions drawn from the resulting analyses.

## 6. Discussion

The extensions introduced in this paper provide the researcher with tools useful in drawing causal inferences. We have formalized the conditions under which signs can be added to the edges of a causal directed acyclic graph and have provided rigorous rules governing their use. We have also argued that a theory based only on average causal effects will not work. The result given in Theorem 3 governing the relationship between monotonic effects and causal effects in the presence of intermediate variables allows the researcher in certain cases to determine a priori whether a particular intervention will on average have a positive or negative effect. The rules governing monotonic effects and covariance given in Theorem 4 may assist the researcher in assessing whether assumptions being made about the causal structure of variables and about monotonic effects are in fact valid. Because of unmeasured variables, control for confounding is often inadequate and various attempts have been made to provide sensitivity bounds for unmeasured confounders. Theorem 5 allows the researcher under certain circumstances to determine the sign of the bias, thereby making clear whether the estimate under consideration is generous or conservative. The directed acyclic graph causal framework has proved to be a useful tool in thinking carefully about questions of estimation and confounding. It is hoped that these contributions will extend yet further the framework's utilization and applicability.

## Appendix A

### A.1. Outline of proof of theorem 1

Let  $P = \{V_1, \dots, V_n\}$  be an ordered list of the nodes on directed paths between  $A$  and  $B$  including  $A$  and  $B$  such that  $A = V_1$  and  $B = V_n$ . It can be shown by induction that by starting with  $n = k$  and for each  $k$  iteratively replacing by their negations the parents of  $V_k$  with negative edges into  $V_k$  suffices for the result.

### A.2. Proof of lemma 3

For  $a \geq a'$  we have

$$\begin{aligned}
& E\{h(Y, A, R)|A = a, R = r\} - E\{h(Y, A, R)|A = a', R = r\} \\
&= \int_{-\infty}^{\infty} h(y, a, r)dF(y|a, r) - \int_{-\infty}^{\infty} h(y, a', r)dF(y|a', r) \\
&= \int_{-\infty}^{\infty} h(y, a, r)d\{F(y|a, r) - F(y|a', r)\} + \int_{-\infty}^{\infty} \{h(y, a, r) - h(y, a', r)\}dF(y|a', r) \\
&= [h(y, a, r)\{F(y|a, r) - F(y|a', r)\}]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \{F(y|a, r) - F(y|a', r)\}dh(y; a, r) \\
&\quad + \int_{-\infty}^{\infty} \{h(y, a, r) - h(y, a', r)\}dF(y|a', r) \\
&\geq \int_{-\infty}^{\infty} \{S(y|a, r) - S(y|a', r)\}dh(y; a, r) + \int_{-\infty}^{\infty} \{h(y, a, r) - h(y, a', r)\}dF(y|a', r).
\end{aligned}$$

This final expression is non-negative since the integrands of both integrals are non-negative for  $a \geq a'$ .

### A.3. Generalization of lemma 4 and outline of proof

*Lemma 4.* Let  $X$  denote some set of non-descendents of  $A$  that blocks all backdoor paths from  $A$  to  $Y$ . Let  $R = (R_1, \dots, R_m)$  denote an ordered list of some set of nodes on directed paths from  $A$  to  $Y$  such that for each  $i$  the backdoor paths from each element of  $R_i, \dots, R_m$  to  $Y$  are blocked by  $R_1, \dots, R_{i-1}, A$  and  $X$ . If all directed paths from  $A$  to  $Y$  are positive except possibly through  $R$  then  $S(y|a, x, r)$  and  $E(y|a, x, r)$  are non-decreasing in  $a$ .

*Outline of proof.* Let  $C$  denote the set of common causes of  $A$  and  $Y$ . Let  $Q$  denote the set of nodes that are ancestors of  $A$  or of  $Y$  but are not descendents of  $A$  and not common causes of  $A$  and  $Y$ . By Theorem 1, considering only those directed paths not blocked by  $R$ , we may replace certain variables by their negations so that all edges on all directed paths from  $A$  to  $Y$  not blocked by  $R$  have positive sign. Let  $V_1 = A$  and  $V_n = Y$  and let  $V_1, \dots, V_n$  be an ordered list of all the nodes on directed paths from  $A$  to  $Y$  such that at least one of the directed paths from each node to  $Y$  is not blocked by  $R$ . Let  $\bar{V}_k = \{V_2, \dots, V_k\}$ . Let  $Q^k, C^k$  and  $R^k$  be the members of  $Q, C$  and  $R$  respectively that are ancestors of  $\bar{V}_k$ . Using several d-separation arguments it can be shown that

$$\begin{aligned}
S(v_k|a, v_2, \dots, v_{k-1}, c, q, r) &= S(v_k|a, v_2, \dots, v_{k-1}, c, q, r^k) \\
&= S(v_k|a, v_2, \dots, v_{k-1}, c^k, q^k, r^k) = S(v_k|pa_{v_k}).
\end{aligned}$$

We can express  $S(y|a, c, q, r) = E\{1(V_n > v)|A, C, Q, R\}$  as

$$E[E[\dots E[E\{1(V_n > v)|A, \bar{V}_{n-1}, C, Q, R]|A, \bar{V}_{n-2}, C, Q, R]|\dots|A, V_2, C, Q, R]|A, C, Q, R]. \quad (1)$$

Now conditional on  $A, \bar{V}_{n-1} \setminus V_i, C, Q, R$  we have that the innermost conditional expectation

$$E\{1(V_n > v)|A, \bar{V}_{n-1}, C, Q, R\}$$

is non-decreasing in  $v_i$  for  $i = 1, \dots, n-1$  since  $V_i$  has either a weak positive monotonic effect or no effect on  $V_n$ . By iteratively applying Lemma 3 to progressively more outward conditional expectation terms in (1), it can be shown that  $S(y|a, c, q, r)$  is non-decreasing in  $a$ . By another d-separation argument it can be shown that  $S(y|a, x, r)$  is non-decreasing in  $a$  and by Lemma 3 it follows that  $E(y|a, x, r)$  is also non-decreasing in  $a$ .

### A.4. Proof of theorem 2

We prove Theorem 2 in the case of weak monotonic effects. The case of monotonic effects is reasonably straightforward and follows by recursive substitution. By the properties of causal directed acyclic graphs, the original graph  $G$  can be marginalized to the causal directed acyclic graph  $H$ . Let  $C$  denote the set of common causes of  $A$  and  $Y$ . Let  $Q$  denote the set of nodes that are ancestors of  $A$  or of  $Y$  but are not descendents of  $A$  and not common causes of  $A$  and  $Y$ .

By Lemma 4 with  $X = \{C, Q\}$ , it follows that  $S(y|a, c, q)$  is non-decreasing in  $a$  and that the edge between  $A$  and  $Y$  is positive on  $H$ . Let  $Q_1$  denote the subset of  $Q$  which are ancestors  $A$ ; let  $Q_2$  denote the subset of  $Q$  which are ancestors  $Y$ . As noted above,  $S(y|a, c, q)$  is non-decreasing in  $a$ . It remains to show that  $S(y|a, c)$  is non-decreasing in  $a$ . Because variables in  $Q_1$  are ancestors of  $Y$  only through  $A$  we have that

$$S(y|a, c, q) = S(y|a, c, q_2).$$

Now variables in  $Q_2$  are neither ancestors nor descendents of  $A$ ; furthermore,  $C$  will contain all common cause of  $A$  and variables  $Q_2$ . From this it follows that  $A$  and  $Q_2$  are d-separated given  $C$ . Thus

$$\begin{aligned} S(y|a, c) &= E[S(y|a, c, Q_2)|a, c] \\ &= E[S(y|a, c, Q_2)|c] = E[S(y|a, c, Q)|c] \end{aligned}$$

and since  $S(y|a, c, Q)$  is non-decreasing in  $a$  for all values of  $y, c, q$  we have that

$$S(y|a, c) = E[S(y|a, c, Q)|c]$$

is non-decreasing in  $a$ .

#### A.5. Proof of theorem 3

Let  $C$  be the parents of  $A$ . By Lemma 4,  $E(Y|A = a, C = c)$  is non-decreasing in  $a$  for all values of  $c$  and by the back-door path criterion we have for  $a_1 \geq a_2$  that

$$\begin{aligned} E(Y_{A=a_1}) &= \int E(Y|A = a_1, C = c)dF(c) \\ &\geq \int E(Y|A = a_2, C = c)dF(c) = E(Y_{A=a_2}). \end{aligned}$$

#### A.5. Proof of theorem 4

We prove the result for positive monotonic association; the result for negative monotonic association follows by replacing the variable  $A$  with  $-A$ . We employ several times through this proof the result that

$$\text{cov}(Y, Z) = \text{cov}\{E(Y|X), E(Z|X)\} + E\{\text{cov}(Y, Z|X)\}.$$

By replacing each  $C_i$  with its negation if necessary we may assume that every path from  $C_i$  to  $A$  and from  $C_i$  to  $Y$  is of positive sign. Let  $C_1, \dots, C_n$  denote an ordered list of the nodes in  $C$ . Applying Theorem 1, we may replace certain nodes with their negations so that all edges on all directed paths from  $A$  to  $Y$  are positive. We may apply the algorithm in the proof of Theorem 1 to the directed paths from  $C_n$  to  $Y$  ignoring all the paths from  $C_n$  to  $Y$  through  $A$  so that by replacing certain variables with their negations all edges on all directed paths from  $C_n$  to  $Y$  not through  $A$  are of positive sign. We may apply the algorithm in the proof of Theorem 1 to the directed paths from  $C_n$  to  $A$  so that by replacing certain variables with their negations all edges on all directed paths from  $C_n$  to  $A$  are of positive sign. Note that the edges on directed paths from  $C_n$  to  $A$  cannot also be the edges on directed paths from  $C_n$  to  $Y$  not through  $A$  because by assumption  $C_n$  has no descendents that are common causes of  $A$  and  $Y$ . For  $C_i$  we may apply the algorithm in the proof of Theorem 1 first to directed paths from  $C_i$  to  $Y$  not through  $A, C_n, \dots, C_{i+1}$  then to directed paths from  $C_i$  to  $A$  not through  $C_n, \dots, C_{i+1}$ , then from  $C_i$  to  $C_n$  not through  $C_{n-1}, \dots, C_{i+1}$ , from  $C_i$  to  $C_j, j > i$  not through  $C_{j-1}, \dots, C_{i+1}$  and finally from  $C_i$  to  $C_{i+1}$ . Applying this argument for each  $i$  we may replace certain variables with their negations so every edge on all directed paths from  $A$  to  $Y$ , from each  $C_i$  to  $A$  and from each  $C_i$  to  $Y$  is positive. By applying an argument similar to that in the proof of Theorem 2 the causal directed acyclic graph can be collapsed to one with only  $A$  and  $Y$  and their common causes  $C$  such that the edge between  $A$  and  $Y$  is positive and such that every edge on all directed paths from each  $C_i$  to  $A$  and from each  $C_i$  to  $Y$  is positive. We then have that

$$\text{cov}(A, Y) = \text{cov}\{E(A|C), E(Y|C)\} + E\{\text{cov}(A, Y|C)\}.$$

We will first show that  $E\{cov(A, Y|C)\}$  is non-negative. We have that

$$\begin{aligned} cov(A, Y|C) &= E[\{A - E(A|C)\}\{Y - E(Y|C)\}|C] \\ &= E(E[\{A - E(A|C)\}\{Y - E(Y|C)\}|A, C]|C) \\ &= E[\{A - E(A|C)\}\{E(Y|A, C) - E(Y|C)\}|C] \\ &= cov\{A, E(Y|A, C)|C\}. \end{aligned}$$

Given  $C$ ,  $E(Y|A, C)$  is a non-decreasing function of  $A$  by Lemma 4. Furthermore, given  $C$ ,  $A$  is a non-decreasing function of  $A$  and thus by Lemma 5 we have that for each  $C$ ,  $cov(A, Y|C) \geq 0$  and so

$$E\{cov(A, Y|C)\} \geq 0.$$

We will now show that  $cov\{E(A|C), E(Y|C)\}$  is non-negative. For each component  $C_i$  of  $C$  we may apply the generalization of Lemma 4 in each case letting  $X$  be the non-descendants in  $C$  of  $C_i$  and letting  $R$  be the descendants in  $C$  of  $C_i$ . We then have that  $E(A|C)$  and  $E(Y|C)$  are non-decreasing in each component of  $C$ . Let  $f(C) = E(A|C)$  and  $g(C) = E(Y|C)$  so that  $cov\{E(A|C), E(Y|C)\} = cov\{f(C), g(C)\}$  where  $f(C)$  and  $g(C)$  are non-decreasing in each component of  $C$ . Let  $S_1$  denote the subset of  $C$  which has no ancestors; let  $S_2$  denote the subset of  $C$  which has no ancestors other than those in  $S_1$ ; let  $S_i$  denote the subset of  $C$  which has no ancestors other than those in  $S_1, \dots, S_{i-1}$  and let  $k$  be such that  $C = S_1, \dots, S_k$ . Note that given  $S_1, \dots, S_{i-1}$  the components of  $S_i$  are independent of one another. We then have that

$$\begin{aligned} cov\{E(A|C), E(Y|C)\} &= cov\{f(C), g(C)\} \\ &= cov[E\{f(C)|S_1\}, E\{g(C)|S_1\}] + E[cov\{f(C), g(C)|S_1\}]. \end{aligned}$$

Applying the conditional covariance result again to the second of these two terms, conditioning on  $S_2$  we have

$$\begin{aligned} cov\{f(C), g(C)\} &= cov[E\{f(C)|S_1\}, E\{g(C)|S_1\}] \\ &\quad + E(cov[E\{f(C)|S_1, S_2\}, E\{g(C)|S_1, S_2\}|S_1]) + E[cov\{f(C), g(C)|S_1, S_2\}]. \end{aligned}$$

And continuing to iteratively apply the conditional covariance result to the final term gives

$$\begin{aligned} cov\{f(C), g(C)\} &= cov[E\{f(C)|S_1\}, E\{g(C)|S_1\}] \tag{2} \\ &\quad + E(cov[E\{f(C)|S_1, S_2\}, E\{g(C)|S_1, S_2\}|S_1]) \\ &\quad + E(cov[E\{f(C)|S_1, S_2, S_3\}, E\{g(C)|S_1, S_2, S_3\}|S_1, S_2]) \\ &\quad + \dots + E(cov[E\{f(C)|S_1, \dots, S_{k-1}\}, E\{g(C)|S_1, \dots, S_{k-1}\}|S_1, \dots, S_{k-2}]) \\ &\quad + E[cov\{f(C), g(C)|S_1, \dots, S_{k-1}\}]. \end{aligned}$$

Consider the  $i$ th term of this expression,

$$E(cov[E\{f(C)|S_1, \dots, S_i\}, E\{g(C)|S_1, \dots, S_i\}|S_1, \dots, S_{i-1}]).$$

Now

$$\begin{aligned} E\{f(C)|S_1, \dots, S_i\} \\ &= E[E[\dots E[E\{f(C)|S_1, \dots, S_{k-1}\}|S_1, \dots, S_{k-2}] \dots |S_1, \dots, S_{i+1}]|S_1, \dots, S_i] \end{aligned}$$

and  $f(C)$  is non-decreasing in  $S_1, \dots, S_{k-1}$  and  $S_k$ . Also,  $S(s_k|s_1, \dots, s_{k-1}) = S(s_k|pa_{S_k})$  is non-decreasing in  $s_1, \dots, s_{k-1}$  since all edges from  $S_1, \dots, S_{k-1}$  to  $S_k$  are of positive sign and so by Lemma 3 we have that  $E\{f(C)|S_1, \dots, S_{k-1}\}$  is non-decreasing in  $S_1, \dots, S_{k-1}$ ; furthermore  $S(s_{k-1}|s_1, \dots, s_{k-2})$  is non-decreasing in  $s_1, \dots, s_{k-2}$  and so  $E[E\{f(C)|S_1, \dots, S_{k-1}\}|S_1, \dots, S_{k-2}]$  is non-decreasing in  $S_1, \dots, S_{k-2}$ . Carrying the argument forward we have that

$$\begin{aligned} E\{f(C)|S_1, \dots, S_i\} \\ &= E[E[\dots E[E\{f(C)|S_1, \dots, S_{k-1}\}|S_1, \dots, S_{k-2}] \dots |S_1, \dots, S_{i+1}]|S_1, \dots, S_i] \end{aligned}$$

is non-decreasing in  $S_1, \dots, S_i$ . Similarly  $E\{g(C)|S_1, \dots, S_i\}$  is non-decreasing in  $S_1, \dots, S_i$  and so conditional on  $S_1, \dots, S_{i-1}$ ,  $E\{f(C)|S_1, \dots, S_i\}$  and  $E\{g(C)|S_1, \dots, S_i\}$  are both non-decreasing functions of the independent random variables  $S_i$  and thus by Lemma 5,

$$\text{cov}[E\{f(C)|S_1, \dots, S_i\}, E\{g(C)|S_1, \dots, S_i\}|S_1, \dots, S_{i-1}] \geq 0$$

and so

$$E(\text{cov}[E\{f(C)|S_1, \dots, S_i\}, E\{g(C)|S_1, \dots, S_i\}|S_1, \dots, S_{i-1}]) \geq 0.$$

Since each term of (2) is non-negative we have that

$$\text{cov}\{E(A|C), E(Y|C)\} = \text{cov}\{f(C), g(C)\} \geq 0.$$

This completes the proof.

#### A.6. Proof of theorem 5

We prove the theorem for the case in which all unblocked backdoor paths are of positive sign; the case in which all unblocked backdoor paths are of negative sign follows by replacing the variable  $A$  with  $1 - A$ . Let  $C$  be the set of ancestors of  $A$  from which there exists a directed path to  $Y$  not through  $A$ . Let  $W$  be the set of ancestors of  $A$  which are not in  $X$  but through which, conditioning on  $X$ , there exists an unblocked backdoor path from  $A$  to  $Y$  a component of which consists of a directed path from the node in  $W$  to  $Y$  i.e.  $W$  is the set of variables in  $C$  that are ancestors of  $Y$  in the graph in which edges out of  $A$  and out of  $X$  are removed. Let  $X' = C \setminus W$ . The estimate of the causal effect of  $A$  on  $Y$  conditioning on  $X$  will be the same as the estimate of the causal effect of  $A$  on  $Y$  conditioning on  $Z = X \cup X'$  because all backdoor paths blocked by  $X' = C \setminus W$  are blocked also by  $X$ . For the true causal effect of  $A$  on  $Y$  it suffices to condition on  $Z \cup W$  since conditioning on  $W$  blocks all backdoor paths not blocked by conditioning on  $Z$ . Let  $W_1, \dots, W_k$  be an ordered list of the variables in  $W$ . For each  $W_i \in W$ , by replacing  $W_i$  with its negation if necessary, we may assume that all directed paths from  $W_i$  to  $A$  are of positive sign and that all directed paths from  $W_i$  to  $Y$  except possibly through  $A$  are of positive sign. The true causal effect on  $Y$  of intervening to set  $A = a$  can be given by

$$\begin{aligned} E(Y_{A=a}) &= \sum_{z,w} E(Y|A = a, Z = z, W = w)pr(W = w, Z = z) \\ &= \sum_z \{ \sum_w E(Y|A = a, Z = z, W = w)pr(W = w|Z = z) \}pr(Z = z). \end{aligned}$$

Let

$$g(a, z, w) = E(Y|A = a, Z = z, W = w)$$

and let  $F_{W|Z=z}$  be the distribution function of  $W$  given  $Z = z$  then

$$E(Y_{A=a}) = \sum_z E_{F_{W|Z=z}} \{g(a, z, W)\}pr(Z = z).$$

The estimate of the causal effect conditioning on  $Z$  only can be given by

$$\begin{aligned} S_1 &= \sum_z E(Y|A = a, Z = z)pr(Z = z) \\ &= \sum_z \{ \sum_w E(Y|A = a, Z = z, W = w)pr(W = w|A = a, Z = z) \}pr(Z = z). \end{aligned}$$

Let  $F_{W|Z=z, A=a}$  be the distribution function of  $W$  given  $Z = z$  and  $A = a$  so that

$$S_1 = \sum_z E_{F_{W|Z=z, A=a}} \{g(a, z, W)\}pr(Z = z).$$

Consider some  $W_i$  in the set  $W$ . Let  $W_i^*$  denote the subset of  $\{W_{i+1}, \dots, W_k\}$  that are descendants of  $W_i$ ; let  $W_i^+ = \{W_{i+1}, \dots, W_k\} \setminus W_i^*$ . We may apply the generalization of Lemma 4 to each  $W_i$ . Let  $Z \cup W_i^+ \cup \{W_1, \dots, W_{i-1}\}$  correspond to  $X$  in the generalization of Lemma 4 and let  $A \cup W_i^*$  correspond to  $R$ . Note that all directed paths from  $W_i$  to  $Y$  not through  $A$  are of positive sign; all backdoor paths from  $W_i$  to  $Y$  are blocked by variables  $\{W_1, \dots, W_{i-1}\}$ ; all backdoor paths from  $A$  to  $Y$  are blocked by variables in  $Z \cup W$ ; for all  $W_j \in W_i^*$  all backdoor paths from  $W_j$  to  $Y$  are blocked by variables in  $\{W_1, \dots, W_{j-1}\}$ . We thus have that

$$g(a, z, w) = E(Y|A = a, Z = z, W = w)$$

is non-decreasing in  $w_i$ . Since this argument holds for all  $i$  we have that  $E[Y|A = a, Z = z, W = w]$  is non-decreasing in  $w$ . By Bayes rule,

$$pr(W = w|A = 1, Z = z) = \frac{pr(A = 1|W = w, Z = z)}{pr(A = 1|Z = z)}pr(W = w|Z = z).$$

By Lemma 4 it follows that

$$pr(A = 1|W = w, Z = z) = E(A|W = w, Z = z)$$

is non-decreasing in each component of  $w$  since all directed paths from any component of  $W$  to  $A$  are of positive sign and since all backdoor paths from  $W_j$  to  $A$  are blocked by variables in  $\{W_1, \dots, W_{j-1}\}$  and  $Z$ . Let

$$v_z(w) = \frac{pr(A = 1|W = w, Z = z)}{pr(A = 1|Z = z)}$$

which we have shown is non-decreasing in  $w$ . The distribution functions  $F_{W|Z=z}$  and  $F_{W|Z=z, A=1}$  are related by

$$pr(W = w|A = 1, Z = z) = v_z(w)pr(W = w|Z = z)$$

and so

$$E_{F_{W|Z=z, A=1}}\{g(a, z, W)\} = E_{F_{W|Z=z}}\{g(a, z, W)v_z(W)\}.$$

Since  $E_{F_{W|Z=z}}\{v_z(W)\} = 1$  we have

$$E_{F_{W|Z=z}}[g(a, z, W)\{v_z(W) - 1\}] = Cov_{F_{W|Z=z}}\{g(a, z, W), v_z(W)\}$$

and since both  $g(a, z, w)$  and  $v_z(w)$  are increasing in  $w$  it follows by an argument very similar to that in Theorem 4 that  $Cov_{F_{W|Z=z}}\{g(a, z, W), v_z(W)\} \geq 0$ . Thus

$$E_{F_{W|Z=z, A=1}}\{g(a, z, W)\} = E_{F_{W|Z=z}}\{g(a, z, W)v_z(W)\} \geq E_{F_{W|Z=z}}\{g(a, z, W)\}$$

and so

$$\begin{aligned} S_1 &= \sum_z E_{F_{W|Z=z, A=1}}\{g(a, z, W)\}pr(Z = z) \\ &\geq \sum_z E_{F_{W|Z=z}}\{g(a, z, W)\}pr(Z = z) = E(Y_{A=1}). \end{aligned}$$

The proof that  $S_0 \leq E(Y_{A=0})$  is similar.

## References

- Archer, N. P. and Wang, S. (1993). Application of the backpropagation neural network algorithm with monotonicity constraints for two-group classification problems. *Decision Sciences* **24**, 60-75.
- Bioch, J. C. (1998). Dualization, decision lists and identification of monotone discrete functions. *Annals of Mathematics and Artificial Intelligence* **13**, 69-91.
- Brumback, B. A., Hernán, M. A., Haneuse, S. J. P. A. and Robins J. M. (2004). Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Statist. Med.* **23**, 749-767.
- Cox, D. R. and Wermuth, N. (2003). A general condition for avoiding effect reversal after marginalization. *J. R. Statist. Soc. B* **65**, 937-941.
- Dawid, A. P. (2000). Causal inference without counterfactuals. *J. Am. Statist. Assoc.* **95**, 407-424.
- Dawid, A. P. (2002). Influence diagrams for causal modelling and inference. *Int. Statist. Rev.* **70**, 161-189.

- Druzdzel, M.J. and Henrion, M. (1993). Efficient reasoning in qualitative probabilistic networks. In *Proceedings of the 11th Annual Conference on Artificial Intelligence*, 548-553. Washington, D.C.
- Duncan, O. D. (1975). *Introduction to Structural Equation Models*. New York: Academic Press.
- Esary, J. D., Proschan, F. and Walkup, D. W. (1967). Association of random variables, with applications. *Ann. Math. Statist.* **38**, 1466-1474.
- Galles, D. and Pearl, J. (1995). Testing identifiability of causal effects. In *Uncertainty in Artificial Intelligence* (ed. P. Besnard and S. Hanks), **11**, 185-195. San Francisco: Morgan Kaufman.
- Geiger, D., Verma, T. S. and Pearl, J. (1990). Identifying independence in bayesian networks. *Networks* **20**, 507-534.
- Geng, Z., Guo, J. and Fung, W. K. (2002). Criteria for confounders in epidemiological studies. *J. R. Statist. Soc. B* **64**, 3-15.
- Haavelmo, T. (1943). The statistical implications of a system of simultaneous equations. *Econometrica* **11**, 1-12.
- Härdle, W., Kerkycharian, G., Picard, D. and Tsybakov, A. (1998). *Wavelets, Approximation, and Statistical Applications*. New York: Springer-Verlag.
- Kuroki, M. and Miyakawa, M. (1999). Identifiability criteria for causal effects of joint interventions. *J. Jap. Statist. Soc.* **29**, 105-117.
- (2003). Covariate selection for estimating the causal effect of control plans using causal diagrams. *J. R. Statist. Soc. B* **65**, 209-222.
- Lauritzen, S. L., Dawid, A. P., Larsen, B. N. and Leimer, H. G. (1990). Independence properties of directed Markov fields. *Networks* **20**, 491-505.
- Lauritzen, S. L. and Richardson, T. S. (2002). Chain graph models and their causal interpretations (with discussion). *J. Roy. Statist. Soc. Ser. B* **64**, 321-361.
- Lin, D. Y., Psaty, B. M. and Kronmal, R. A. (1998). Assessing the sensitivity of regression results to unmeasured confounding in observational studies. *Biometrics* **54**, 948-963.
- Manski, C. F. (1997). Monotone treatment response. *Econometrica* **65**, 1311-1334.
- Pearl, J. (1993). Comment: Graphical models, causality, and intervention. *Statist. Sci.* **8**, 266-269.
- (1995). Causal diagrams for empirical research. *Biometrika* **82**, 669-688.
- (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.
- Pearl, J. and Robins, J. M. (1995). Probabilistic evaluation of sequential plans from causal models with hidden variables. In *Uncertainty in Artificial Intelligence* (ed. P. Besnard and S. Hanks), **11**, 444-453. San Francisco: Morgan Kaufman.
- Potharst, R. and Feelders, A. (2002). Classification trees for problems with monotonicity constraints. *Special Interest Group on Knowledge Discovery and Data Mining Explorations* **4**, 1-10.

- Robins, J. M. (1986). A new approach to causal inference in mortality studies with sustained exposure period - application to control of the healthy worker survivor effect. *Mathematical Modelling* **7**, 1393-1512.
- (1987). Addendum to a new approach to causal inference in mortality studies with sustained exposure period - application to control of the healthy worker survivor effect. *Computers and Mathematics with Applications* **14**, 923-945.
- (1995). Discussion of "Causal diagrams for empirical research" by J. Pearl. *Biometrika* **82**, 695-698.
- (1997). Causal inference from complex longitudinal data. In *Latent Variable Modeling and Applications to Causality* (ed. M. Berkane), Lecture Notes in Statistics **120**, 69-117. New York: Springer Verlag.
- (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In *Highly Structured Stochastic Systems* (eds. P. Green, N.L. Hjort, and S. Richardson), 70-81. New York: Oxford University Press.
- Rosenbaum P. R. and Rubin, D. B. (1983a). The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**, 41-55.
- (1983b). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *J. R. Statist. Soc. B* **45**, 212-218.
- Spirites, P., Glymour, C. and Scheines, R. (1993). *Causation, Prediction and Search*. New York: Springer-Verlag.
- Tian, J. and Pearl, J. (2002). On the identification of causal effects. In *Proceedings of the Eighteenth National Conference on Artificial Intelligence*, 567-573. Menlo Park: AAAI Press/The MIT Press.
- van der Gaag, L. C., Bodlaender, H. L. and Fielders, A. (2004). Monotonicity in Bayesian networks. In *Uncertainty in Artificial Intelligence* (ed. Chickering, M. and Halpern, J.), **20**, 569-576. San Francisco: Morgan Kaufman. Menlo Park: AAAI Press/The MIT Press.
- VanderWeele, T. J. and Robins, J. M. (2007). Directed acyclic graphs, sufficient causes and the properties of conditioning on a common effect. *Am. J. Epidemiol.*, in press.
- Verma, T. and Pearl, J. (1988). Causal networks: Semantics and expressiveness. In *Proceedings of the 4th Workshop on Uncertainty in Artificial Intelligence*, 352-359. Reprinted in *Uncertainty in Artificial Intelligence* (ed. R. Shachter, T.S. Levitt, and L.N. Kanal), **4**, 69-76. Amsterdam: Elsevier.
- Wellman, M. P. (1990). Fundamental concepts of qualitative probabilistic networks. *Artificial Intelligence* **44**, 257-303.
- Wright, S. (1921). Correlation and causation. *J. Agric. Res.* **20**, 557-585.