The New Challenge: From a Century of Statistics to an Age of Causation

J. Pearl
Computer Science Department
University of California, Los Angeles
Los Angeles, CA 90095-1596

Abstract

Some of the main users of statistical methods – economists, social scientists, and epidemiologists – are discovering that their fields rest not on statistical but on causal foundations. The blurring of these foundations over the years follows from the lack of mathematical notation capable of distinguishing causal from equational relationships. By providing formal and natural explication of such relations, graphical methods have the potential to revolutionize how statistics is used in knowledge-rich applications. Statisticians, in response, are beginning to realize that causality is not a metaphysical dead-end but a meaningful concept with clear mathematical underpinning. The paper surveys these developments and outlines future challenges.

1 A Century of Denial

Francis Galton's discovery of correlation, at the end of the nineteenth century [Galton, 1888], dazzled one of his students, Karl Pearson, generally considered the founder of modern statistics. The year 1911 saw publication of the third edition of Pearson's *The Grammar of Science*, which contained a new chapter titled "Contingency and correlation - the insufficiency of causation." This is how Pearson introduces the new topic: "Beyond such discarded fundamentals as 'matter' and 'force' lies still another fetish amidst the inscrutable arcana of modern science, namely, the category of cause and effect" [Pearson, 1911, p. *iv*]. And what does Pearson substitute for the archaic notion of causation? Correlations and contingency tables!! He states (ibid, p. 159),

> Such a table is termed a *contingency* table, and the ultimate scientific statement of description of the relation between two things can always be thrown back upon such a contingency table. . . . Once the reader realizes the nature of such a table, he will have grasped the essence of the conception of association between cause and effect.

Thus, Pearson categorically denies the need for a concept of causal relation independent of or beyond correlation. He held this view throughout his life and, accordingly, did not mention causation in any of his technical papers. His objection to animistic concepts such as "will" and "force" was so fierce and his rejection of determinism so absolute that he exterminated causation from statistics before it had a chance to take root.

Pearson's crusade influenced markedly the direction of statistical research and education in the twentieth century, also known as "The Statistical Century." The *Encyclopedia of Statistical Science* [Kotz and Johnson, 1982], for example, devotes 12 pages to correlation but only 2 pages to causation, and one of those pages is spent demonstrating that "correlation does not imply causation." Given the dearth of doctoral theses, research papers, and textbook pages on causation, Pearson apparently still rules statistics.

Modern statisticians acknowledge the stalemate over causality. Philip Dawid states, "Causal inference is one of the most important, most subtle, and most neglected of all the problems of statistics" [Dawid, 1979]. Terry Speed declares; "Considerations of causality should be treated as they have always been treated in statistics: preferably not at all, but if necessary, then with very great care" [Speed, 1990]. David Cox and Nanny Wermuth, in a book published just a few months ago, explain, "We did not in this book use the words *causal* or *causality*.... Our reason for caution is that it is rare that firm conclusions about causality can be drawn from one study" [Cox and Wermuth, 1996]. This caution about and avoidance of causality has influenced many fields that look to statistics for guidance, especially economics, the social sciences, and the health sciences. This statement from one leading social scientist is typical: "It would be very healthy if more researchers abandon thinking of and using terms such as cause and effect" [Muthen, 1987].

How can we explain why statistics, the field that gave the world powerful ideas such as the testing of hypotheses and the design of experiment abandoned causation so easily and so early? One obvious explanation is that causation is much harder to measure than correlation. Correlations can be estimated directly in a single uncontrolled study, while causal conclusions require either controlled experiments or causal assumptions of some kind, and these are hard to come by in a single study. But this explanation is simplistic. Statisticians are not easily deterred by difficulties or by the need to conduct several studies, if necessary; and children manage to learn cause-effect relations without running controlled experiments.

The answer lies deeper, and it has to do with the official language of statistics, namely, the language of probability. This should not come as a surprise to most of us, since the word *cause* is not in the vocabulary of probability theory. We cannot express in the language of probabilities the sentence "Mud does not cause rain"; all we can say is that the two events are mutually correlated, or dependent – meaning that if we find one, we can expect the other. Naturally, if we lack a language to express a certain concept explicitly, we cannot expect to develop scientific activity around that concept. "Every science is only so far exact as it knows how to express one thing by one sign," said Augustus de Morgan in 1858, explaining why logic made no progress from the time of Aristotle until the introduction of logical notation. In statistics, a scientific handling of causality would require a language in which the causal relationship "Mud does not cause rain" receives symbolic representation that is clearly distinct from "Mud is independent of rain." Such a language, to the best of my knowledge, has not so far become part of standard statistical research.

2 Researchers in Search of a Language

Two languages for causality have in fact been proposed: path analysis or structural equation modelling (SEM) [Wright, 1921; Haavelmo, 1943], and Neyman-Rubin's potential-response model [Neyman, 1923; Rubin, 1974]. The former has been adopted by economists and social scientists [Goldberger, 1972; Duncan, 1975], while the latter has been advocated by a small but iconoclastic group of statisticians [Rubin, 1974; Robins, 1986; Holland, 1988] who refuse to sanction the official casting of causality out of the province of statistics. Unfortunately, neither of these languages has become part of standard statistical research – the structural equation framework because it has been greatly misused and inadequately formalized [Freedman, 1987], and the potential-response framework because it has been only partially formalized and, more significant, because it rests on an esoteric and seemingly metaphysical vocabulary of counterfactual variables that bears no apparent connection to ordinary understanding of cause-effect processes.

Currently, SEM is used by many and understood by few, while potential-response models are understood by few and used by even fewer. The explanation for this state of affairs may serve as a classic illustration of the immense importance of mathematical notation in the development of the sciences. A brief sketch of the SEM episode follows.

SEM was developed by geneticists [Wright, 1921] and economists [Haavelmo, 1943] so that cause-effect information could be combined with statistical data to answer policy-related questions. Yet current SEM practitioners are constantly tormented by the question, "Under what conditions can we give causal interpretation to identified structural coefficients?" Sewall Wright and Trygve Haavelmo would have answered simply, "Always!" According to the founding fathers of SEM, the conditions that make a set of equations *structural* and a specific equation $y = \beta x + \epsilon$ *identified* are precisely those that make the causal connection between X and Y have no other value but β. Amazingly, this basic understanding of SEM has all but disappeared from the literature on SEM, in both econometrics and the social sciences.

Most SEM researchers today are of the opinion that extra ingredients are necessary for the conclusions of a SEM study to turn into legitimate causal claims. Kenneth Bollen [1989], for example, states that a condition called "isolation" or "pseudo-isolation" is necessary.[1] Bullock, Harlow, and Mulaik [1994] reiterate the necessity of isolation and lament: "confusion has grown concerning the correct use of and the conclusions that can be legitimately drawn from these [SEM] methodologies." Social scientists are not alone in this; the econometric literature has no less difficulty dealing with the causal reading of structural parameters. Ed Leamer [1985] observes, "It is my surprising conclusion that economists know very well what they mean when they use the words 'exogenous,' 'structural,' and 'causal,' yet no textbook author has written adequate definitions." Attempts to overcome this formal deficiency with statistical vo-

[1] Bollen [1989, p. 44] defines pseudo-isolation as the orthogonality condition $cov(x, \epsilon) = 0$, where ϵ is the error term in the equation $y = \beta x + \epsilon$. This condition is neither necessary (as seen, e.g., in the analysis of instrumental variables [Bollen, 1989, pp. 409–413], and in Figure 6 (c, e) of [Pearl, 1995a]) nor sufficient (e.g., [Cartwright, 1995b, p. 50]) unless causal meaning is already attached to β.

cabulary have led to complex definitions of causality [Sims, 1972] and exogeneity [Engle *et al.*, 1983] that exacted a heavy toll before their limitations were brought to light (see [Leamer, 1985; Aldrich, 1993]).

Current difficulties with the causal reading of econometric equations are captured by Steven LeRoy [1995]: "It is a commonplace of elementary instruction in economics that endogenous variables are not generally causally ordered, implying that the question 'What is the effect of y_1 on y_2' where y_1 and y_2 are endogenous variables is generally meaningless." According to LeRoy's recent proposal, causal relationships cannot be attributed to any variable whose causes have separate influence on the effect variable, thus denying causal reading to most of the structural parameters that economists labor to estimate and ruling out most policy variables in economics [Balke and Pearl, 1995].

Nancy Cartwright, a renowned philosopher of economics, addresses these difficulties by initiating a renewed attack on the tormenting question: "*Why* can we assume that we can read off causes, including causal order, from the parameters in equations whose exogenous variables are uncorrelated?" [Cartwright, 1995b]. Like the founders of SEM, Wright and Haavelmo, Cartwright recognizes that causes cannot be derived from statistical or functional relationships alone and that causal assumptions are prerequisite for validating any causal conclusion. Unlike them, however, she launches an all-out search for the assumptions that would endow the parameter β in a regression equation $y = \beta x + \epsilon$ with a legitimate causal meaning and labors to prove that the assumptions she proposes are indeed sufficient. What is revealing in Cartwright's analysis is that she does not consider the answer Haavelmo would have provided (one that applies to models of any size and shape, including models with correlated exogenous variables): the assumptions needed for drawing causes from parameters are encoded in the *syntax* of the equations and can be read off the associated graph as easily as a shopping list[2]; they need not be searched for elsewhere, nor do they require specialized proofs of sufficiency.

Cartwright's analysis reflects an alarming tendency among economists and social scientists to view a structural model as an algebraic object that carries functional and statistical assumptions but is void of causal content.[3] Perhaps the boldest expression of this trend has recently been voiced by Holland [1995]: "I am speaking, of course, about the equation: $\{y = a + bx + \epsilon\}$. What does it mean? The only meaning I have ever determined for such an equation is that it is a shorthand way of describing the conditional distribution of $\{y\}$ given $\{x\}$."[4] A causality-free conception of SEM may explain both Cartwright's search for causal assumptions outside the model and the urge of SEM researchers to fortify the equations with extra conditions (e.g., isolation) or ban the natural causal readings of the equations [LeRoy, 1995].

The founders of SEM expressed no such trepidation. Wright [1923] did not hesitate to declare that "prior knowledge of the causal relations is assumed as prerequisite" in the theory of path coefficients, and Haavelmo [1943] explicitly interpreted each structural equation as a statement about a hypothetical controlled experiment. One wonders, therefore, what has happened to SEM over the past 50 years, and why the basic teachings of Wright and Haavelmo have been forgotten.

I believe that the causal content of SEM has been allowed to gradually escape the consciousness of SEM practitioners mainly for the following reasons:

1. SEM practitioners have sought to gain respectability for SEM by keeping causal assumptions implicit, since statisticians, the arbiters of respectability, abhor such assumptions because they are not directly testable.

2. The algebraic, graph-less language that has dominated SEM research lacks the notational facility needed for making causal assumptions, as distinct from statistical assumptions, explicit. By failing to equip causal relations with distinct mathematical notation, the founding fathers in fact committed the causal foundation of SEM to oblivion. Their disciples today are seeking foundational answers elsewhere.

Let me elaborate on this last point. The founders of SEM understood quite well that the equality sign in structural models conveys the asymmetrical relation "is determined by," and hence it behaves more like an assignment symbol (:=) in programming languages than like an ordinary algebraic equality. However, perhaps for

[2] Specifically, if G is the graph associated with a causal model that renders a certain parameter identifiable, then the assumptions sufficient for authenticating the causal reading of that parameter can be read off G as follows: Every missing arrow, say between X and Y, represents the assumption that X has no causal effect on Y once we intervene and hold the parents of Y fixed. Every missing bi-directed link between X and Y represents the assumption that there are no common causes for X and Y, except those shown in G.

[3] Notable exceptions are [Leamer, 1985] and [Hoover, 1995, pages 75–90].

[4] Holland's interpretation stands at variance with the structural reading of the equation above [Haavelmo, 1943], which is "In an ideal experiment where we control X to x and any other set Z of variables (not containing X or Y) to z, Y is independent of z and is given by $a + bx + \epsilon$" [Pearl, 1995a, p. 704].

reasons of mathematical purity (to avoid the appearance of syntax sensitivity), they refrained from introducing a symbol to represent this asymmetry.

According to Roy Epstein [1987], Wright once gave a seminar on path coefficients to the Cowles Commission (the breeding ground for SEM) in the 1940s, but neither side saw particular merit in the other's methods. Why? After all, a diagram is nothing but a set of nonparametric structural equations in which, to avoid confustion, the equality signs are replaced with arrows.

My explanation is that early econometricians were extremely careful mathematicians; they thought they could keep the mathematics in purely equational-statistical form and just reason about structure in their heads. Indeed, they managed to do so surprisingly well, because they were truly remarkable individuals and *could* do it in their heads. The consequences began to surface in the early 1980s, when their disciples began to mistake the equality sign for an algebraic equality and, suddenly, the "so-called disturbance terms" did not make any sense at all [Richard, 1980]. We are living with the sad end of this tale: by failing to cast their insights in mathematical notation, the founders of SEM brought about the current difficulties surrounding the interpretation of structural equations, as summarized by Holland's "What does it mean?"[5]

3 Graphs as a Mathematical Language

Certain developments in the past decade promise to bring causality back into the mainstream of scientific investigation. These developments involve an improved understanding of the relationships between graphs and probabilities, on one hand, and between graphs and causality, on the other. The fundamental change of the past decade is the emergence of graphs as a mathematical language for causality. By mathematical language, I do not mean simply a heuristic mnemonic device for displaying "deeper" mathematical relationships but quite the opposite: graphs emerge as the fundamental notational system for concepts and relationships that are not easily expressed in any mathematical language (e.g., equations or probabilities) other than graphs. Additionally, graphs can serve both as models for determining the truth of causal utterances and as a symbolic machinery for deriving such truths from other causal premises [Galles and Pearl, 1996].

A concrete example will demonstrate the power and potential of the graphical language. One of the most frustrating issues in causal analysis has been the problem of *covariate selection*, for example, determining whether one can add a variate z to a regression equation without biasing the result. More generally, whenever we try to evaluate the effect of one factor (X) on another (Y), we wonder whether we should adjust our measurements for possible variations in some other variable, Z, sometimes called a *covariate*, *concomitant*, or *confounder*. Adjustment amounts to partitioning the population into groups that are homogeneous relative to Z, assessing the effect of X on Y in each homogeneous group, and, finally, averaging the results.

The elusive nature of such an adjustment was recognized as early as 1899, when Pearson discovered what in modern terms is called *Simpson's paradox*, namely, that any statistical relationship between two variables may be reversed or negated by including additional factors in the analysis. For example, we may find that students who smoke obtain higher grades than those who do not smoke; but, adjusting for age, smokers obtain lower grades than nonsmokers in every age group; but, further adjusting for family income, smokers obtain higher grades than nonsmokers in every income-age group; and so on.[6]

Despite a century of analysis, Simpson's reversal phenomenon continues to "trap the unwary" [Dawid, 1979], and the main question – whether an adjustment for a given covariate Z is appropriate in any given study – continues to be decided informally, on a case-by-case basis, with the decision resting on folklore and intuition rather than on hard mathematics. The standard statistical literature is remarkably silent on this issue and, aside from the common advice that one should not adjust for a covariate that is affected by the putative cause (X), it provides no guidelines as to what covariates would be admissible for adjustment and what assumptions would be needed for making this determination formally.[7]

In the potential-response framework, a criterion called "ignorability" has been advanced [Rosenbaum and Rubin, 1983], which reads: Z is an ad-

[5]The teachings of current economists and philosophers who understand the role of causality in SEM, among them Leamer [1985], Woodward [1995], Cartwright [1995b], Hoover [1995], and Goldberger [1991], are in danger of meeting a similar fate, unless their ideas are cast into mathematical symbols.

[6]The classical case demonstrating Simpson's reversal is the study of Berkeley's alleged sex bias in graduate admission [Bickel *et al.*, 1975], where, overall, data show a higher rate of admission among male applicants but, when broken down by departments, data show a slight bias toward female applicants.

[7]This advice, which rests on the causal relationship "not affected by" is, to the best of my knowledge, the *only* nonstatistical notion that has managed to find a place in statistics textbooks. The advice is, of course, necessary, but it is not sufficient. The other common guideline, that X should not precede Z [Shafer, 1996, p. 326], is neither necessary nor sufficient.

missible covariate relative to the effect of X on Y if, for every x, the value that Y would obtain had X been x is conditionally independent of X, given Z. Needless to say, such a criterion merely paraphrases the problem in the language of counterfactuals without providing a working test for covariate selection. Since counterfactuals are not observable, and judgments about conditional independence of counterfactuals are not readily assertable from ordinary understanding of causal processes, ignorability has remained a theoretical construct that has had only minor impact on practice. Practicing epidemiologists, for example, well apprised of ignorability analysis via the admirable papers of Robins [1986] and Greenland and Robins [1986], are still debating the meaning of "confounding" and often adjust for the wrong sets of covariates [Weinberg, 1993]. Social scientists, likewise, despite a penetrating ignorability analysis of the Lord paradox (a version of Simpson's paradox) by Holland and Rubin [1983], are still struggling with various manifestations of this paradox in psychometric research [Wainer, 1991].

In comparison, formulating the adjustment problem in the language of graphs has immediately yielded a general solution to the problem that is both natural and formal. The solution method invites the investigator to express causal knowledge (read: assumptions) in meaningful qualitative terms by using arrows among quantities of interest, and, once the graph is completed, a simple procedure decides whether a proposed adjustment is appropriate relative to the quantity under evaluation.

For example, the procedure described in the following five steps determines whether a set of variables Z should be adjusted for when we we wish to evaluate the total effect of X on Y. The assumptions encoded in the initial graph were explicated in footnote 2, and figures illustrating the result of each step are given in Appendix I.

Procedure:[8]

Input: Directed acyclic graph in which three subsets of nodes are marked X, Y, and Z.

Output: A decision whether the effect of X on Y can be determined by adjusting for Z.

Step 1. Exit with failure if any node in Z is a descendant of X.

Step 2. (simplification) Simplify the diagram by eliminating all nodes (and their incident edges) that are not ancestors of either X or Y or Z.

Step 3. (triangulation) Add an undirected edge between any two ancestors of Z which share a common child.

Step 4. (pruning) Eliminate all arrows emanating from X.

Step 5. (symmetrization) Strip the arrows from all directed edges.

Step 6. (test) If, in the resulting undirected graph, Z intercepts all paths between X and Y, then Z is an appropriate covariate for statistical adjustment. Else, Z should not be adjusted for.

When failure occurs in Step 1, it does not mean that the measurement of Z cannot be useful for estimating the effect of X on Y; nonstandard adjustments might then be used instead of the standard method of partitioning into groups homogeneous relative to Z (see [Galles and Pearl, 1995]). Finally, if the objective of the study is to evaluate the "direct," rather than the "total," effect of X on Y, as is the case in the Berkeley example, then other graphical procedures are available to determine the appropriate adjustment (see [Pearl and Robins, 1995]).[9]

The example above is not an isolated case for which clarity and precision are gained through the use of graphical methods. In fact, the conceptual basis for SEM achieves a new level of mathematical precision and clarity through graphs. What makes a set of equations "structural," what assumptions should be ascertained by the authors of such equations, and what policy claims are advertised by a given set of structural equations are some of the concerns not addressed formally in the economics literature [Leamer, 1985] that now receive simple and mathematically precise answers.

It turns out that the assumptions encoded in a causal graph are also sufficient for defining other notions that economists have found difficult to interpret – for example, defining when a variable is exogenous, when a variable is an "instrument," and what those "so-called disturbance terms" are.[10] The common definition for exo-

[8]This procedure is an adaptation of the back-door criterion [Pearl, 1993; Pearl, 1995a] using the triangulation test [Lauritzen *et al.*, 1990] of *d*-separation [Pearl, 1988]. An equivalent procedure can be obtained from Theorem 7.1 of Spirtes et al. [1993].

[9]Procedures for proper evaluation of the direct effect of X on Y should embody the requirement that other factors (of Y) should be "held constant" by external means, as distinct from the routine procedure of "adjusting" for those factors.

[10]Readers will recognize the connection between exogeneity and the problem of covariate selection; a variable X is exogenous relative to Y if the effect of X on Y can be determined by regressing Y on X or, in other words, if the empty set of covariates $Z = \{0\}$ is admissible according to the procedure above.

geneity, according to which X is exogenous for Y whenever it is "independent of the random terms of the system" [Dhrymes, 1970, p. 169] is ambiguous, because (1) the random terms are not fully defined [Leamer, 1985] and (2) in the case where the equation for Y contains variables other than X, one must specify which random terms are to be considered. Such difficulties prompted Engle et al. [1983] to seek new definitions of exogeneity, outside of the structural equation framework; however, the definition they finally adopted (i.e., "superexogeneity") turned out merely a complicated disguise of the one they abandoned [Aldrich, 1993].

The potential-response model of Rubin, Holland, and Robins also receives foundational support from the graphical representation. The unit-response function $Y(x, u)$, which is taken as a primitive in the potential-response framework (read: the value that Y would have obtained in unit u had X been x), can now be given a more mathematical interpretation (read: the solution for Y of a given set of simultaneous equations, which is obtained after deleting the equation for X and substituting the conditions $U = u$ and $X = x$). Accordingly, rules of inference that in the potential-response framework must be taken as axioms turn into theorems in the graphical framework, the validity of which rest on the equation-deletion semantics of $Y(x, u)$. Robins' rule of consistency [Robins, 1986]

$$X = x \implies Y(x) = Y$$

is an example of such an axiom/theorem.

How do scientists predict the outcome of one experiment from the results of other experiments run under totally different conditions? Such transfers of experimental knowledge, although essential to scientific progress, involve inferences that cannot easily be formalized in the standard languages of logic, physics, or probability because these influences require a symbolic distinction between manipulative phrases, such as "holding Z fixed," and observational phrases, such as "conditioning on Z." The standard algebras, including the algebra of equations, Boolean algebra, and probability calculus, are all geared to serve observational but not manipulative sentences.

Graphs fill this linguistic gap. They provide both semantics and axiomatic characterization of manipulative statements of the form "Changing X will not affect Y if we hold Z constant," and also serve as "theorem provers" to facilitate the derivation of such sentences from other sentences [Galles and Pearl, 1996].

4 The Challenge

Recent progress in graphical methods and nonparametric structural modeling has rendered causal analysis amiable to ordinary statistical techniques and accessible to rank-and-file researchers. Investigators can now articulate qualitative causal assumptions in a friendly formal language, combine these assumptions with statistical data, and derive new causal conclusions with mathematical precision. Simple methods are now available for solving the following problems:

1. Deriving algebraic expressions for causal effect estimands, both total and direct [Pearl, 1995a].

2. Selecting measurements (covariates or confounders) to obtain unbiased estimates of treatment effects from observational studies (provided certain causal connections can be assumed nonexistent) (see Section 3)

3. Predicting (or bounding) treatment effectiveness from trials with imperfect compliance [Balke and Pearl, 1993; Pearl, 1995b; Chickering and Pearl, 1996].

4. Estimating (or bounding) counterfactual probabilities (e.g., John was treated and died, but would he have survived had he not been treated?) [Balke and Pearl, 1994].

Commenting on the state of logic prior to the advent of Boolean algebra, Augustus de Morgan [1864] observed,

> Every science that has *thriven* has thriven upon its own symbols: logic, the only science which is admitted to have made no improvements in century after century, is the only one which has *grown no symbols.*

Throughout the twentieth century, the study of causality in statistics has been conducted within the confines of probability calculus; it has grown no symbols and has not thriven either. Given the dazzling progress of logic after the advent of Boolean notation, one cannot help but hope that similarly spectacular changes will attend causal modeling once graphical notation is accepted.

Acknowledgments

This research was partially supported by Air Force grant #AFOSR/F496209410173, NSF grant #IRI-9420306, and Rockwell/Northrop Micro grant #94-100.

References

[Aldrich, 1993] J. Aldrich. Cowles' exogeneity and core exogeneity. Technical Report Discussion Paper 9308, Department of Economics, University of Southampton, U.K., 1993.

[Balke and Pearl, 1993] A. Balke and J. Pearl. Nonparametric bounds on causal effects from partial compliance data. Technical Report R-199, Computer Science Department, University of California, Los Angeles, September 1993. To appear in *JASA*.

[Balke and Pearl, 1994] A. Balke and J. Pearl. Counterfactual probabilities: Computational methods, bounds, and applications. In R. Lopez de Mantaras and D. Poole, editors, *Uncertainty in Artificial Intelligence 10*, pages 46–54. Morgan Kaufmann, San Mateo, CA, 1994.

[Balke and Pearl, 1995] A. Balke and J. Pearl. Counterfactuals and policy analysis in structural models. In P. Besnard and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, pages 11–18. Morgan Kaufmann, San Francisco, 1995.

[Bickel *et al.*, 1975] P.J. Bickel, E.A. Hammel, and J.W. O'Connell. Sex bias in graduate admissions: Data from berkeley. *Science*, 187:398–404, 1975.

[Bollen, 1989] K.A. Bollen. Causality and causal models. In *Structural Equations with Latent Variables*, chapter 3. John Wiley and Sons, New York, NY, 1989.

[Bullock *et al.*, 1994] H.E. Bullock, L.L. Harlow, and S.A. Mulaik. Causation issues in structural equation modeling research. *Structural Equation Modeling*, 1(3):253–267, 1994.

[Cartwright, 1995a] N. Cartwright. Causal structures in econometrics. In D. Little, editor, *On the Reliability of Economic Models*, pages ??? – ??? Kluwer Academic Publishers, Boston, 1995.

[Cartwright, 1995b] N. Cartwright. Probabilities and experiments. *Journal of Econometrics*, 67:47–59, 1995.

[Chickering and Pearl, 1996] D.M. Chickering and J. Pearl. A clinician's apprentice for analyzing non-compliance. pages 1269–1276, Portland, OR, August 1996. Proceedings of the National Conference on Artificial Intelligence (AAAI-96).

[Cox and Wermuth, 1996] D.R. Cox and N. Wermuth. *Multivariate Dependencies – Models, Analysis and Interpretation*. Chapman and Hall, London, 1996.

[Dawid, 1979] A.P. Dawid. Conditional independence in statistical theory. *Journal of the Royal Statistical Society, Series A*, 41:1–31, 1979.

[de Morgan, 1864] A. de Morgan. On the syllogism. *Transactions of the Cambridge Philosophical Society*, 10:173–230, 1864. (Read 8 Feb 1958).

[Dhrymes, 1970] P.J. Dhrymes. *Econometrics*. Springer-Verlag, New York, 1970.

[Engle *et al.*, 1983] R.F. Engle, D.F. Hendry, and J.F. Richard. Exogeneity. *Econometrica*, 51:277–304, March 1983.

[Epstein, 1987] R.J. Epstein. *A History of Econometrics*. Elsevier Science Publishing, New York, 1987.

[Freedman, 1987] D. Freedman. As others see us: A case study in path analysis (with discussion). *Journal of Educational Statistics*, 12:101–223, 1987.

[Galles and Pearl, 1995] D. Galles and J. Pearl. Testing identifiability of causal effects. In P. Besnard and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, pages 185–195. Morgan Kaufmann, San Francisco, 1995.

[Galles and Pearl, 1996] D. Galles and J. Pearl. Axioms of causal relevance. Technical Report R-240, Department of Computer Science, University of California, Los Angeles, 1996. To appear in *Artificial Intelligence Journal*.

[Galton, 1888] F. Galton. Co-relations and their measurement, chiefly from anthropological data. *Proceedings of the Royal Society of London*, 45:135–145, 1888.

[Goldberger, 1991] A.S. Goldberger. *A Course of Econometrics*. Harvard University Press, Cambridge, MA, 1991.

[Greenland and Robins, 1986] S. Greenland and J. Robins. Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology*, 15:413–419, 1986.

[Haavelmo, 1943] T. Haavelmo. The statistical implications of a system of simultaneous equations. *Econometrica*, 11:1–12, 1943.

[Holland and Rubin, 1983] P.W. Holland and D.B. Rubin. On Lord's paradox. In H. Wainer and S. Messick, editors, *Principals of Modern Psychological Measurement*, pages ??? – ??? Earlbaum, Hillsdale, NJ, 1983.

[Holland, 1995] P.W. Holland. Some reflections on Freedman's critiques. *Foundations of Science*, 1:50–57, 1995.

[Hoover, 1995] K. Hoover. Comments on Cartwright and Woodward: Causation, estimation, and statistics. In D. Little, editor, *On the Reliability of Economic Models*, pages ??? – ??? Kluwer Academic Publishers, Boston, 1995.

[Kotz and Johnson, 1982] Samuel Kotz and Norman L. Johnson, editors. *Encyclopedia of Statistical Sciences*, New York, 1982. John Wiley and Sons, Inc.

[Lauritzen *et al.*, 1990] S .L. Lauritzen, A. P. Dawid, B. N. Larsen, and H. G. Leimer. Independence properties of directed Markov fields. *Networks*, 20:491–505, 1990.

[Leamer, 1985] E.E. Leamer. Vector autoregressions for causal inference? *Carnegie-Rochester Conference Series on Public Policy*, 22:255–304, 1985.

[LeRoy, 1995] Stephen F. LeRoy. Causal orderings. In Kevin D. Hoover, editor, *Macroeconometrics: Developments, Tensions, Prospects*, pages 211–228. Kluwer Academic Publishers, Boston, 1995.

[Muthen, 1987] B. Muthen. Response to Freedman's "As others see us: A case study in path analysis". *Journal of Educational Statistics*, 12:168–175, 1987.

[Pearl and Robins, 1995] J. Pearl and J.M. Robins. Probabilistic evaluation of sequential plans from causal models with hidden variables. In P. Besnard and S. Hanks, editors, *Uncertainty in Artificial Intelligence 11*, pages 444–453. Morgan Kaufmann, San Francisco, 1995.

[Pearl, 1988] J. Pearl. *Probabilistic Reasoning in Intelligence Systems.* Morgan Kaufmann, San Mateo, CA, 1988. Revised 2nd printing, 1992.

[Pearl, 1993] J. Pearl. Comment: Graphical models, causality and intervention. *Statistical Science*, 8:266–269, August 1993.

[Pearl, 1995a] J. Pearl. Causal diagrams for experimental research. *Biometrika*, 82:669–710, December 1995.

[Pearl, 1995b] J. Pearl. Causal inference from indirect experiments. *Artificial Intelligence in Medicine*, 7:561–582, 1995.

[Pearson, 1911] K. Pearson. *Grammar of Science,* 3rd ed. A. and C. Black Publishers, ??????, 1911.

[Richard, 1980] J.F. Richard. Models with several regimes and changes in exogeneity. *Review of Economic Studies*, 47:1–20, 1980.

[Robins, 1986] J.M. Robins. A new approach to causal inference in mortality studies with a sustained exposure period - applications to control of the healthy workers survivor effect. *Mathematical Modeling*, 7:1393–1512?, 1986.

[Rosenbaum and Rubin, 1983] P. Rosenbaum and D. Rubin. The central role of propensity score in observational studies for causal effects. *Biometrica*, 70:41–55, 1983.

[Shafer, 1996] G. Shafer. *The Art of Causal Conjecture.* MIT Press, Cambridge, MA, 1996.

[Sims, 1972] C. Sims. Money, income, and causality. *American Economic Review*, 62:540–552, 1972.

[Speed, 1990] T.P. Speed. Complexity calibration and causality in influence diagrams. In R.M. Oliver and J.Q. Smith, editors, *Influence Diagrams, Belief Nets and Decision Analysis*, page 58. John Wiley and Sons, New York, 1990.

[Spirtes *et al.*, 1993] P. Spirtes, C. Glymour, and R. Schienes. *Causation, Prediction, and Search.* Springer-Verlag, New York, 1993.

[Wainer, 1991] H. Wainer. Adjusting for differential base-rates: Lord's paradox again. *Psychological Bulletin*, 109:147–151, 1991.

[Weinberg, 1993] C.R. Weinberg. Toward a clearer definition of confounding. *American Journal of Epidemiology*, 137:1–8, 1993.

[Woodward, 1995] J. Woodward. Causation and explanation in econometrics. In D. Little, editor, *On the Reliability of Economic Models*, pages ??? – ??? Kluwer Academic Publishers, Boston, 1995.

[Wright, 1921] S. Wright. Correlation and causation. *Journal of Agricultural Research*, 20:557–585, 1921.

[Wright, 1923] S. Wright. The theory of path coefficients: A reply to Niles' criticism. *Genetics*, 8:239–255, 1923.

Papers by J. Pearl are available at http://singapore.cs.ucla.edu/jp_home.htm

(Appendix I on following page.)

Appendix I

Figure 1: The graphical solution of the covariate-selection problem.