Some Statistical Aspects of Causality

D.R. Cox and Nanny Wermuth

A general review of approaches to causality is given from a statistical perspective. Three broad notions are distinguished. In the final part of the paper the challenges of reaching potentially causal representations are outlined for a study of some German political and social attitudes.

Generalities

There is a long history in the philosophical literature of discussions of causality. It typically regards a cause as necessary and sufficient for an effect: all children of divorced parents have behavioural problems and all children with behavioural problems have divorced parents. In virtually all sociological contexts, however, the concern is with multiple causes, even if one is predominant. Thus explicit or implicit statistical considerations are inescapable to evaluate empirical evidence. Even within the statistical view of causality there are a number of different formulations. Here we review three. In the final section of the paper a specific illustration is sketched. References to work quoted by authors only are in Holland (1986).

Notions of Causality

Causality as Stable Association

Suppose that there is clear evidence that two features of the individuals (people, communities, households, etc.) under investigation are associated, a possible cause C and a possible response R. For instance, individuals with high values of C may tend to have high values of R and vice versa. Thus C and R might be test scores for individuals at a given age in arithmetic and language or they may be the unemployment rate and level of crime in a community. What might it mean to conclude that C is a cause of a response R?

Symmetric and Directed Relations. Association is a symmetric relation between two or possibly more features, but causality is asymmetric. That is, if C is associated with R then R is associated with C, but if C is a cause of R then R is not a cause of C. Thus, given any two features C and R, we need to distinguish the possibilities where

1. C and R are to be treated on an equal footing and dealt with symmetrically in any interpretation;
2. One of the variables, say C, is to be regarded as explanatory to the other variable, R, regarded as a response. Then, if there is a relation, it is regarded asymmetrically.

The distinction here is not about statistical significance but rather is concerned with substantive interpretation.

Graphical Representation. A useful graphical representation shows two variables X_1 and X_2, regarded on an equal footing, if associated, as connected by an undirected edge, whereas two variables such that C is explanatory of R, if connected, are done so by a directed edge: see Figures 1*a* and 1*b*.

There are two possible bases for the distinction between explanatory and response variables and thus for using a directed edge. One is that features referring to an earlier time-point are explanatory to features referring to a later time-point. Thus aspects of previous education may be possible explanatory variables for subsequent career performance. In such situations the relevant time is not the time when the observation is made but the time to

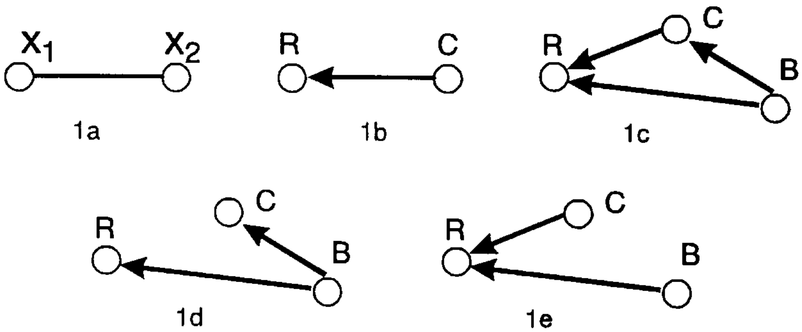

Figure 1.
(a) Undirected edge between two variables X_1, X_2 *on an equal footing*
(b) Directed edge between explanatory variable C *and response variable* R
(c) General dependence of response R *on* B, C
(d) Special situation with $R \perp\!\!\!\perp C \mid B$
(e) Special situation with $B \perp\!\!\!\perp C$ *corresponding in particular to randomization of* C

which the features refer although, of course, observations recorded retrospectively are especially subject to recall biases. The second is a subject-matter working hypothesis based for example on theory or on empirical data from other kinds of investigation. Suppose that cross-sectional data are collected on current salary and on attitudes to various social and political issues. It might then be reasonable provisionally to regard income as explanatory of attitudes, of course only as part of a more complex network of relationships.

In summary, the first step towards causality is to require good reasons for regarding C as explanatory of R as a response and that any notion of causal connection between C and R is that C is a cause of R and not the other way round.

Common Explanatory Variables. Next consider the possibility of one or more common explanatory variables. For this, suppose that a background variable B is potentially explanatory of C and hence also of R. There are a number of possibilities of which the most general is shown in Figure 1*c* with directed edges from B to C, from C to R and also directly from B to R. On the other hand, if the relation were that represented schematically in Figure 1*d*, the only dependence between C and R is that induced by their both depending on B. Then C and R are said to be conditionally independent given B, sometimes conveniently written $R \perp\!\!\!\perp C \mid B$. There is no direct path from C to R that does not pass via B. Such relations are typically assessed empirically by some form of regression analysis. In such a situation one would not regard C as a cause of R even though in an analysis without the background variable B there is a statistical dependence between C and R.

This discussion leads to one definition used in the literature of C being a cause of R, namely that there is a dependence between C and R and that the sign of that dependence is unaltered whatever variables $B_1, B_2, \ldots$ themselves explanatory to C are considered simultaneously with C as possible sources of dependence. This definition has a long history but is best articulated by I. J. Good and P. Suppes. A corresponding notion for time series is due to N. Wiener and C.W. Granger. This definition underlies much empirical statistical analysis in so far as it aims to achieve causal explanation.

The definition entertains all possible alternative explanatory variables. In any particular study one can at best check that the measured background variables B do not account for the dependence between C and R. The possibility that the dependence could be explained by variables explanatory to C that have not been measured, i.e. by so-called unobserved confounders, is less likely the larger the apparent effect and can be discounted only by general plausibility arguments about the field in question. Sensitivity analysis may be helpful as it involves calculating what the properties of an unobserved confounder would have to be to explain away the dependence in question. When the empirical dependence found is

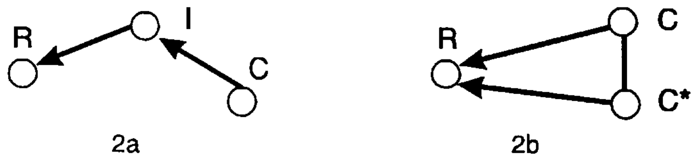

Figure 2.
(a) Intermediate variable I *accounting for overall effect of* C *after ignoring* I; R $\perp\!\!\!\perp$ C | I
(b) Associated variables C,C* *on an equal footing and both explanatory to* R

very strong it may be possible to argue that it is implausible that there is an unmeasured variable itself with a strong enough dependence to explain away completely the observed effect. For further details see Rosenbaum (1995).

Role of Randomization. Although physical randomization by the investigator is rarely possible in sociological investigations, it is conceptually important to consider briefly its consequences for interpretation. In this case in the scheme sketched in Figure 1*e* there can be no edge between the *B*s and *C*, since such dependence would be contrary to randomization, i.e. to each individual under study being equally likely to receive each treatment possibility. In this situation an apparent dependence between *C* and *R* cannot be explained by a background variable as in Figure 1*d*. It is in this sense that causality can be inferred from randomized experiments and not from observational studies as sometimes stated, especially in the statistical literature. While, other things being equal, randomized experiments are greatly to be preferred to observational studies, difficulties of interpretation, sometimes serious ones, remain. The most important are possible interactive effects, especially with unobserved explanatory variables, and unanticipated future interventions in the system under study that remain unnoticed when the final response is recorded.

Intermediate Variables. In the above discussion the variables *B* have been supposed to be explanatory of *C* and hence of *R*. For judging a possible causal effect of *C* it would be wrong to consider in the same way variables intermediate between *C* and *R*, i.e. variables *I* that are responses to *C* and explanatory of *R*. Although they are valuable in clarifying the nature of any indirect path between *C* and *R*, the use of *I* as an explanatory variable in a regression analysis of *R* on *C* would not be enough in assessing whether such a path exists. If *R* is independent of *C* given an intermediate variable *I*, but dependent on *I*, then *C* may still have caused *I* and *I* may be a cause of *R*.

For instance suppose that *C* represents an aspect of primary school education and *I* some feature of secondary education, which in turn affects some aspect of employment; see Figure 2*a*. Does the aspect of primary education cause a change in *R*? If *R* is conditionally independent of *C* given *I* it would be reasonable to say that primary education does cause a change in *R* and that this change appears to be explained via what happens in secondary education.

Explanatory Variables on an Equal Footing. An even more delicate situation arises with variables C^* on an equal footing with the variable *C* whose causal status is under consideration; see Figure 2*b*. If the role of *C* is essentially the same whether or not C^* is conditioned, i.e. whether or not C^* is included in the regression equation, there is no problem, at least at a qualitative level. On the other hand it is relatively common to find clear dependence on (C, C^*) as a pair, but that either variable on its own is sufficient to explain the dependence. There are then broadly three routes to interpretation:

1. To regard (C, C^*) collectively as the possibly causal variables;
2. To present at least two possibilities for interpretation, one based on *C* and one on C^*;
3. To obtain further information clarifying the relation between *C* and C^*, establishing for instance that C^* is explanatory of *C* and that the appropriate interpretation is to fix C^* when analysing variations in *R*.

For example, suppose that *C* and C^* are respectively measures of educational performance in arithmetic and language of a child both measured at the same age and that the response is some adult feature. Then the third possibility is inapplicable; the first possibility is to regard the variables as a two-dimensional measure of educational performance and to abandon, at least temporarily, any notion of separating the role of arithmetic and language. By taking the second route we would recognize two alternative simple explanations consistent with the data, one based on arithmetic and one on language. This is typically studied via the careful examination of empirical regression equations, for example for binary R via logistic regression.

Causality as the Effect of Intervention

Counterfactuals. The concept of causality discussed above is important and is connected with the approach adopted in many empirical statistical studies. It does not, however, directly capture a stronger interpretation of the word causal. This is connected with the idea of hypothetical intervention or modification. Suppose for simplicity of exposition that C takes just two possible forms, to be called presence and absence. Thus presence might be the implementation of some programme of intervention, and absence a suitable control state. For monotone relations one may say that the presence of C causes an increase in the response R if an individual with C present tends to have a higher R than that same individual would have had if C had been absent, other things being equal.

Slightly more explicitly let B denote all variables possibly explanatory to C and suppose that there are no variables C^* to be considered on an equal footing to C. Consider for each individual two possible values of R, R_{pres} and R_{abs}, that would arise as C takes on its two possible values present and absent and B is fixed. Then the presence of C causes, say, an increase in R if R_{pres} is in some sense systematically greater than R_{abs}. The notion of other things being equal is captured by holding B fixed.

This notion is in part a translation of J. Neyman's and R. A. Fisher's work on the design of experiments into a more general setting, including an observational one. It is connected with the work of H. A. Simon and has been systematically studied and very fruitfully applied by D. B. Rubin.

For a given individual only one of R_{pres} and R_{abs} can be observed, corresponding to the value of C actually holding for that individual. The other value of R is a so-called counterfactual whose introduction is, however, useful to capture the notion hinted at above of a deeper meaning of causality.

Formalizing Differences in Counterfactuals. The simplest and least demanding relation between the two values of R is that over some population of individuals under study the average of R_{pres} exceeds that of R_{abs}. This is a notion of an average effect and is testable empirically in favourable circumstances. A much stronger requirement is that the required inequality holds for every individual in the population of concern. Stronger still is the requirement that the difference between the two values of R is the same for all individuals, i.e. that for all individuals

$$R_{\text{pres}} - R_{\text{abs}} = \Delta.$$

In the language of the theory of the design of experiments this is called the assumption of unit-treatment additivity.

Now these last two assumptions are clearly not directly testable and can be objected to on that account. The assumptions are indirectly testable, to a limited extent at least. If the individuals are divided into groups, for example on the basis of one or more of the background variables B, the assumptions imply that for each individual observed that the difference between the two levels of R has the same sign in the first case and the second case. However, overlooking the fact that a possibly causal variable C has a very different effect on different individuals can have severe consequences.

Intrinsic Variables or Attributes. There is an important restriction implicit in this discussion. It has to be meaningful in the context in question to suppose that a potential cause C for an individual might have been different from how it in fact is. This is relevant only to variables that appear solely as explanatory variables. For example they may be variables measured at some base-line, i.e. at entry into a study. Purely explanatory variables can be divided into intrinsic variables or attributes, sometimes also called structural variables, which are essentially defining characteristics of the individual, and potential explanatory variables which might play the role of C in the present discussion. Intrinsic variables should not be regarded as potentially causal in the present sense. For example the gender of an individual is in most contexts an intrinsic characteristic. The question what would R have been for this woman had she been a man other things being held fixed is in many, although not quite all, contexts meaningless.

Variables to be Held Fixed. Finally, care is essential in defining what is to be held fixed under hypothetical changes in C. In terms of statistical analysis this is the issue of what other variables should be included as explanatory variables in the regression equation for R in addition to C itself. Certainly responses I to C are not fixed. Variables, B, explanatory of C are held fixed. There is an essential ambiguity for variables C^* on an equal footing with C. To distinguish changing C to a given level with certain other features held fixed from the probabilistic notion of conditioning Pearl (1995) has introduced the terminology of *setting* C to its required level; see also Pearl (2000) and Lauritzen (2000) for further discussion.

Causality as Explanation of a Process

There is a third notion of causality that is in some ways most in line with normal scientific usage. In this context causality implies that there is some understanding, albeit provisional, of the process that leads from C to R. This understanding typically comes from theory, or often from knowledge at a hierarchical level lower than the data under immediate analysis. Sometimes, it may be possible to represent such a process by a graph without directed cycles and to visualize the causal effect by the tracing of paths from C to R via variables I intermediate between C and R. Thus the effect of interventions at a community level might be related to ideas of individual psychology.

This last notion of causality as concerned with generating processes is to be contrasted with the second view of causality as concerned with the effects of intervention and with the first view of causality as stable statistical dependence. These views are complementary not conflicting. Goldthorpe (this volume) has argued for this third view of causality as the appropriate one for sociology with explanation via rational choice theory as an important route for interpretation.

To be satisfactory there needs to be evidence, typically arising from studies of different kinds, that such generating processes are not merely hypothesized. Causality is not to be established by merely calling a statistical model causal.

Explanations of phenomena in terms of underlying processes are inevitably provisional. Nevertheless they are the cornerstone of the natural sciences. We suggest that statistical analysis should aim towards establishing processes that are potentially causal.

Special Issues

Interaction Involving a Potentially Causal Variable

We now turn to the issue of interactions with a potentially causal variable. The graphical representations used above to show the structure of various kinds of dependency and independency holding between a set of variables have the limitation, in the form used here, that they do not represent interaction, in particular that an effect of C may be systematically different for different groups of individuals. For example if B is an intrinsic feature such as gender, we consider whether the effect of C is different for men and for women. In particular, if the effects of C are in opposite directions for different levels of B we say there is a qualitative interaction, a possibility of special importance for interpretation.

Note especially that even when C represents a randomized treatment which is automatically decoupled from preceding possibly unobserved variables B the possibility of serious interactions with B cannot in general be ignored.

Viewed slightly differently, absence of interaction is important not only in simplifying interpretation but also in enhancing generalizability and specificity. That is, an effect that has been shown to have no serious interaction with a range of potential variables is more likely to be reproduced in some new situation and more likely to have a stable subject-matter interpretation.

Unwanted Unobserved Intermediate Variable

Consider further the role of variables I referring to time points after the implementation of C. A subject-matter distinction can be drawn between, on the one hand, intermediate variables that are responses to C and that are explanatory to R and are part of some natural process, and, on the other hand, interventions into the system that may depend on C and which may be explanatory of R but which in some sense are unwanted or inappropriate for interpretation. Thus in studying the effect of modifications in inner-city housing policies on satisfaction it may be necessary to take account of interventions other than those which are the immediate object of study. Another example is in evaluations of study programmes whenever students in only one of the programmes receive unplanned intensive encouragement during the evaluation period.

Aggregation

So far little has been said about the choice of observational units for study. At a fundamental research level it may be wise to choose individuals showing the effects of interest in their simplest and most striking form. More generally, however, the choice has to be considered at two levels. There is the level at which ultimate interpretation and action is required and the level at which careful observation of largely decoupled individuals is available. For example, a criminologist comparing different sentencing or policing policies is interested in individual offenders but may be able to observe only different communities or policing areas. A nutritional epidemiologist comparing different diets is interested in individual people but may have to rely, in part at least, on consumption and mortality data from whole countries. The assumption that a dependence established on an aggregate scale, for example at a country level, has a similar interpretation at a small-scale level, for example for individual persons, involves the assumption that there are no confounders B at the person level that would account for the apparent dependency. This will typically be very hard or even impossible to check awith any accuracy from country-level data.

Bradford Hill's Conditions

The above discussion implicitly emphasizes that, while causal understanding is the aim of perhaps nearly all research work, a cautious approach is essential, especially but not only in observational studies. The most widely quoted conditions tending to make a causal interpretation more likely are those of Hill (1965) put forward in connection with the interpretation of epidemiological studies. Hill emphasized their tentative character. For a critical discussion of these conditions, see Rothman and Greenland (1998). Because they are usually mentioned in an epidemiological context we reproduce them in outline in a slightly revised version (Cox and Wermuth, 1996; sect. 8.7).

According to these conditions a dependency is more likely to be causal

- if an *a priori* subject-matter explanation of it is available;
- if a convincing subject-matter explanation is found retrospectively although such is typically less convincing than an *a priori* explanation;
- if the effect is a large one, it then being less likely that there is an alternative explanation via an unmeasured confounder;
- if the dependency has a natural monotonic relation with levels of the explanatory variable in question;
- if the effect is found repeatedly in independent studies especially if these are of somewhat different form;
- if there is no major interaction with intrinsic features;
- if the dependence is the consequence of a massive intervention in the system.

The bacteriologist Koch gave conditions for inferring causality when the potential cause can be applied, withdrawn, and reapplied in a relatively controlled way and the pattern of response observed. Similar ideas were used in psychology following early experiments by Pavlov with conditioning stimuli and the extinction of responses after withdrawing the stimulus.

A Sociological Case Study

Causal formulations are common in social science contexts whenever empirical research is planned to study explanations for the development of opinions, attitudes, and judgements or behaviour. In these cases those statistical models and analyses are especially helpful which permit the exclusion of prior expectations on a developmental process since they provide interpretations that are compatible with causal explanations. We give an example of deriving a graphical Markov model that concerns the development of attitudes towards interventions by the state. In the case of increasing unemployment should the state intervene? For the general well-being of society should the state provide a social safety net?

The data used are from the German General Social Surveys (Central Archive for Empirical Social Research, 1985; 1992). They are cross-sectional surveys which started in 1980 and are typically carried out every second year. A special additional survey was conducted in 1991 in both East and West Germany. Since questions on attitudes towards state interventions and on their possible determinants were asked only in 1991 and in 1984, i.e. before unification, only West Germans are included in the analyses. And, since we study the risk of exclusion from the workforce as one of the possible determinants the data are restricted to individuals who are between 18 and 65 years old.

Figure 3 shows an initial ordering of the variables with four boxes indicating four possible stages of development, ranging from given background or context variables on the far right to the response variable of primary interest on the far left: attitude towards state interventions, Y, a sum score formed from two questions with high values denoting agreement with interventions. Relations among context variables are known from many studies; they are taken to be fixed in the present analysis, i.e. they are not analysed here, they are conditioned on. Two context variables are taken as quantitative measurements: age, W, and education of parents, V, (0=none , 1=one, or 2=both have at least 12 years of formal schooling); the three other context variables are binary: own education, B, (1=at least 12 years of formal schooling), marital status, C, (1=married), and gender, D, (1=female), all taken to have values zero and one.

Risk of exclusion from the workforce, A, (1=yes) is a rough indicator constructed from information available in the data on unemployment, on the qualifications necessary to proceed to university, and on completed vocational training. It is listed alone in one of the middle boxes, indicating that it is an intermediate variable, regarded both as a possible response to each of the context variables (to its right) and as a possible explanatory variable for the remaining variables (to its left). Two attitudes are

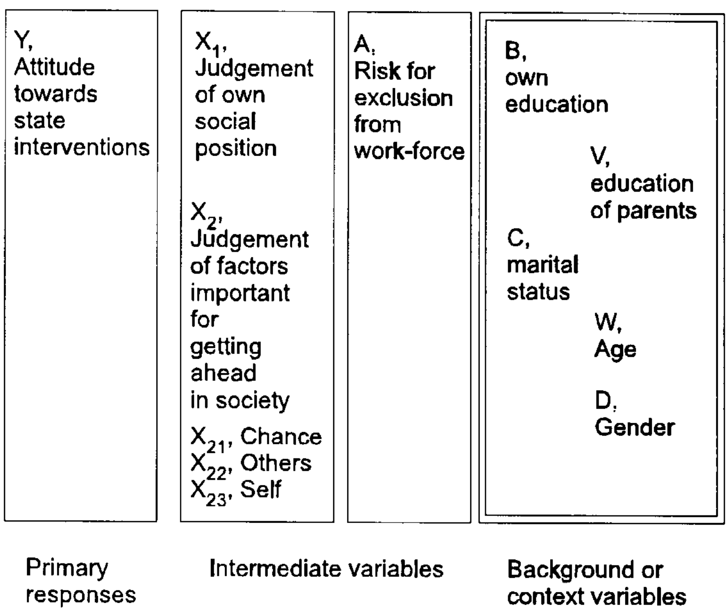

Figure 3. *An initial ordering of the variables of possible determinants of attitudes towards interventions by the state*

expected to be of a more stable kind: judgement of own social position, X_1, (1=lower class to 5= upper class) and judgement on which factors are important for getting ahead in society, X_2. There are three sepecial aspects of the latter: chance, X_{21}, others, X_{22}, (wealthy family, general connections (in 1991), protection (in 1984), political connections, money, and opportunism), or oneself, X_{23}, (own education, hard work, initiative, and intelligence). For each of these measurements a high value indicates strong agreement that the aspect is important.

Since we do not consider any two of the four judgements to relate as response and explanatory variables they are treated as being on an equal footing, i.e. they are listed in the same box, so that some joint response model will be used for their analysis. Jointly they are intermediate variables, i.e. possibly explanatory for a respondent's attitude towards state interventions and possible responses to the risk and context variables.

The ordering of the boxes provides a plan for the type of analyses to be carried out: for instance a linear regression for the attitudes towards state intervention on all other variables, a logistic regression for the risk on all the context variables, and a multivariate regression for the joint responses given the risk and the context variables. For the purpose of checking for independences we replace the multivariate regression of the X-variables by univariate analyses for each response taken in turn with all (or a subset of) the variables to the right as explanatory variables.

Standard checks for non-linear relations and interactions did not reveal very strong effects of either type. Univariate distributions for the raw data and for reduced sets of data with complete observations on all 11 variables (693 cases in 1991 and 1727 cases in 1984) appeared essentially unchanged. The main changes from 1984 to 1991 in univariate distributions are an increase in the percentage of persons with at least 12 years of formal schooling, from 21 to 27 per cent, and a decrease in the percentage of persons at risk of exclusion from the workforce, from 25 to 20 per cent.

Figure 4 summarizes the analyses conducted separately for the two years by showing an arrow for each variable which turned out to be an important explanatory factor in either year. For instance the contributions of each variable with a missing arrow to the response of primary interest would have been very minor, i.e. corresponding to a t-value smaller in absolute value than 1.5 if included next as an explanatory variable in the regresssion. A dashed line indicates that a substantial correlation among two joint responses remained after accounting for all the effects of variables being important for either one. None of the considered context variables was explanatory for the getting-ahead scales. This can be interpreted positively as the lack of item and scale bias.

By using known relations among the context variables in addition to the regression results, parts of which are shown in Tables 1 to 3, one of the pathways of development in Figure 4 (Y, X_1, A, B, V) can be interpreted as follows: if a respondent's parents have less formal education it is more likely that his or her own formal education will be shorter. The less formal education one has the higher is the risk of exclusion from the workforce. A high risk of exclusion from the workforce makes it more likely that the respondent's own social position is judged to be low. Persons who judge their own social position to be low tend to agree strongly with the type of state interventions considered here. Additional factors for predicting high agreement with state interventions are the

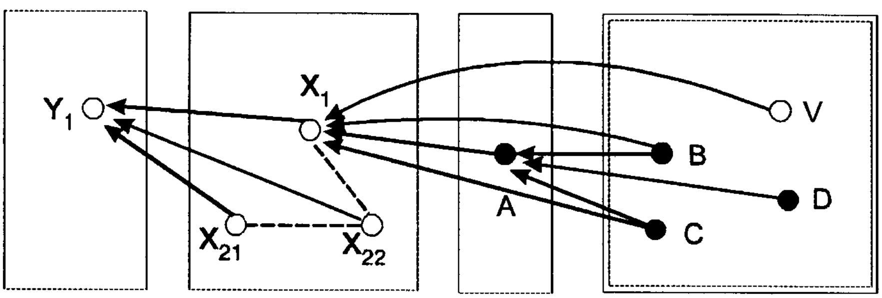

Figure 4. *A chain graph developed as a result of analyses*

Table 1. *Regression coefficients for response* Y, *attitude towards state interventions*

Explanatory variable	1984	1991
X_1, social position	−0.35	−0.46
X_2, getting ahead		
X_{21}, chance	0.11	0.15
X_{22}, others	0.07	0.04
constant	5.99	6.35

Table 2. *Regression coefficients for response* X_1, *judgement of own social position*

Explanatory variable	1984	1991
A, risk for low social position	−0.17	−0.18
B, own education	0.39	0.40
C, married	0.04	0.15
V, education of parents	0.22	0.17
constant	2.68	2.67

Table 3: Regression coefficients for response A, *risk of exclusion from workforce*

Explanatory variable	1984	1991
B, own education	−1.16	−0.97
C, married	−0.31	−0.73
D, gender	0.74	0.23
constant	−1.11	−0.86

beliefs that chance or others determine what is important to get ahead in society.

Probably the most remarkable finding of this study is the strong agreement in qualitative conclusions in the years 1984 and 1991. The regression coefficients shown for directly important factors determining attitudes towards state interventions (in Table 1) are essentially identical. For the judgement of one's own social position (in Table 2) the general level and the effects of the predictors of risk for exclusion from the workforce and own education are unchanged over the years, while the importance of parents' education has decreased and the effect of marital status is strong only in 1991. Although the overall risk of exclusion from the workforce has decreased in Germany from 1984 to 1991 the direction of dependencies has remained unchanged for own education, marital status, and gender. The risk is higher for less formal schooling, for females, and for the unmarried. Only the relative importance of gender has diminished; that of marital status has increased over the years. Thus, the type of dependencies are replicated except when political changes can explain changes in dependencies.

The interpretation in terms of tracing paths is close to the original suggestions of Sewall Wright for path analyses in linear systems. Graphical Markov models are one extension of path analysis in which responses, intermediate and explanatory variables may be quantitative or categorical variables and in which joint instead of only single responses may be modelled. Linear structural equation models are a different extension of path analysis. The main differences from a chain graph model such as that in Figure 4 are twofold: the way in which categorical response variables are modelled, and the interpretation of parameters. In chain graphs every edge in the graph, missing or present, can be interpreted as a particular conditional relationship, so that the vanishing of an edge means a conditional independence statement. The interpretation of parameters in a linear structural equation model may have to be derived from scratch if it does not coincide with a chain graph model.

In summary, the analysis sketched here aims to find pathways of dependence from the baseline variables through the intermediate variables to the final response variable. While, of course, conclusions from a single observational study are inevitably provisional, the approach brings us a step closer to the notions of causality discussed in the first part of the paper. The analysis gives more insight than, for example, only direct regression analysis of the final response on all available potentially explanatory variables.

Some References on Causality

The extensive and growing literature on statistical aspects of causality is best approached via the discussion paper of by Holland (1986); see also Cox and Wermuth (1996; sect. 8.7). For general issues about observational studies, see Cochran (1965) and Rosenbaum (1995). For a philosophical perspective, see Simon (1972) and Cartwright (1989). For an interventionist view, see Rubin (1974) and for a more formal approach still from a social science viewpoint Sobel (1995). For a systematic account based on directed acyclic graphs, see Pearl (1995, 2000) and for the general connections with graph theory see Lauritzen (2000). For an approach based on a complete specification of all independencies between a set of variables followed by a computer-generated listing of all directed acyclic graphs consistent with those independencies, see Spirtes *et al*. (1993). The use of counterfactuals is criticized by Dawid (2000). Be aware that many rather different interpretations of causality are involved in these discussions.

References

Cartwright, N. (1989) *Nature's Capacities and their Measurement*. Clarendon Press, Oxford.

Central Archive for Empirical Social Research (Zentralarchiv für Empirische Sozialforschung) and ZUMA (1985) *ALLBUS Basisumfrage 1984*. Codebuch ZA-Nr.1340. Zentralarchiv, Köln.

Central Archive for Empirical Social Research (Zentralarchiv für Empirische Sozialforschung) & ZUMA (1992) *ALLBUS Basisumfrage 1991*. Codebuch ZA-Nr.1990. Zentralarchiv, Köln.

Cochran, W. G. (1965) The planning of observational studies of human populations (with discussion). *Journal of the Royal Statistical Society* A **128**, 234–265.

Cox, D. R., Wermuth, N. (1996) *Multivariate Dependencies: Models, Analysis and Interpretation*. Chapman and Hall, London.

Dawid, A. P. (2000) Causality without counterfactuals (with discussion). *Journal of the American Statistical Association* **95**, 407–448.

Hill, A. B. (1965) The environment and disease: association or causation. *Proceedings of the Royal Society of Medicine*, **58**, 295–300.

Holland, P. (1986) Statistics and causal inference (with discussion). *Journal of the American Statistical Association*, **81**, 945–970.

Lauritzen, S. L. (2000) Causal inference from graphical models. In Klüppelberg, C., Barndorff-Nielsen, O. E. and Cox, D. R. (eds.) *Complex Stochastic Systems*. Chapman and Hall/CRC, London.

Pearl, J. (1995) Causal diagrams for empirical research (with discussion). *Biometrika*, **82**, 669–710.

Pearl, J. (2000) *Causality: Models, Reasoning and Inference*. Cambridge University Press, Cambridge.

Rosenbaum, P. R. (1995) *Observational Studies*. Springer, New York.

Rothman, K. J., Greenland, S. (1998) (eds.) *Modern Epidemiology*, 2nd edn. Raven-Lippincott, Philadelphia.

Rubin, D. B. (1974) Estimating causal effects of treatment in randomized and nonrandomized studies. *Journal of Educational Studies*, **66**, 688–701.

Simon, H. A. (1972) Causation. In Kruskal, W. H., Tanur and J. M. (eds.) *Encyclopedia of Statistics*, i. Free Press, New York, pp. 35–41.

Sobel, M. E. (1995) Causal inference in the social and behavioral sciences. In Arminger, G., Clogg, C. C. and Sobel, M. E. (eds.) *Handbook of Statistical Modeling for the Social and Behavioral Sciences*. Plenum, New York.

Spirtes, P., Glymour, C., and Scheines, R. (1993) *Causation, Prediction and Search*. Springer, New York.

Wermuth, N. (1998) Graphical Markov models. In Kotz, S., Read, C. and Banks, D. (eds) *Encyclopedia of Statistical Sciences*. Wiley, New York: 2nd Update vol., pp. 284–300.

Author's Address

Professor D. R. Cox, Nuffield College, Oxford OX1 1NF, U.K. Tel.: +44 (0) 1865 278690; fax: +44 (0) 1865 278621; e-mail: david.cox@nuf.ox.ac.uk.