I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2013/43

Confounding and Control in a Multivariate System
An issue in causal attribution

RUSSO, F., MOUCHART, M. AND G. WUNSCH

Confounding and Control in a Multivariate System

An issue in causal attribution

Federica RUSSO [a], Michel MOUCHART [b] and Guillaume WUNSCH[c]

[a] *Dipartimento di Studi Umanistici, Università degli Studi di Ferrara, Italy*

[b] *CORE and ISBA, Université catholique de Louvain, Belgium*

[c] *Demography, Université catholique de Louvain & Royal Academy of Sciences, Belgium*

December 3, 2013

Abstract

It is widely agreed that, in establishing whether variable X causes variable Y, a third variable Z may confound the relation and thus hinder causal assessment. The solution developed within the 'traditional' framework is to control for any third variable, susceptible of confounding the relation between X and Y. This paper examines complex systems of variables, characterised by multiple causes and multiple effects. The paper advances the view that in such contexts confounding is a moot issue, under a suitable specification of the causal model. When networks of causal relations are considered, possible confounders are included in the appropriate causal paths from the causes to the outcome. The challenge for the model builder then amounts to developing a structural model that specifies the role of variables in each path, rather than just controlling for possible confounders.

Keywords: Causality, Confounding, Control, Structural Modelling.

Corresponding Author: Federica Russo, Dipartimento di Studi Umanistici, Università degli Studi di Ferrara, Italy, email rssfrc@unife.it

Contents

1 Introduction

The problem of confounding has been recognized by methodologists and philosophers alike as one of the major threats in establishing genuine causal relations. Given a correlation between two variables, say A and B, it is possible that a third variable C confounds a putative causal relation between A and B insofar as it is a common cause of A and B. For example, one can observe a positive association between income and health: the higher the income the better the health. Though a higher income might indeed lead to better health care, the association might also be due to the level of education of the individuals concerned: higher education often leads to a higher income and is favorable to better information on health behaviors and to better health practices. In this elementary example, education confounds the relation between income and health; it is therefore desirable to control for education in order to isolate the proper possible impact of income on health.

More generally, controlling for a confounding variable intends to evaluate the specific impact of a particular putative cause, acting in a complex system, on a particular variable of interest, *i.e.* a system with at least three variables, of which one for the cause and another one for the effect, the third variable being a possible confounder of the cause and effect relation. In our example, the three variables would be income (putative cause), health (outcome or effect), and education (confounder).

The 'third variable' problem is of course a real one, once we consider simple causal relations between two variables. In practice, however, one is confronted with a situation where a relation between a putative cause and an effect is immersed in a complex directed (or ordered) network of relationships among a larger number of variables, which includes possible confounders. In this paper, we examine the issue of confounding and control in complex systems from the perspective of structural modeling. Structural models specify a suitable mechanism of the data generating process. Therefore, the task is to identify the directed network of various cause-effect relationships, and not just a single 'one-cause one-effect' relation. The modeller has to order those variables as correctly as possible, in order to recursively represent the underlying data generating process (DGP). In this perspective, the issue of controlling for a possible confounder amounts to giving confounding variables the correct place within the web of causal relations, which often includes multiple mechanisms leading to the effect. This is the purpose of the structural modelling approach presented later in this paper. In this perspective, confounding is a moot issue because confounders are taken into account in the postulated mechanism.

The paper is organized as follows. In section 2 we present confounding and control as traditionally considered. In this paper we frame the discussion within observational studies carried out *e.g.* in demography, but the concepts of confounding and control apply, of course, to experimental settings too. In section 3 we present the issue of confounding and control in directed networks in more formal terms. In section 4, we propose structural modelling as an alternative framework when considering the issue of confounding and control. We conclude that, while traditional approaches

have considered confounding variables as hindering causal assessment, in a structural modelling framework confounders simply belong to the set of variables to take into account in the *whole* network of relations representing a mechanism or data generating process.

2 Confounding and control: the 'third-variable' problem

In this section, we consider the issue of confounding and control in a traditional setting, that is where the modeller seeks to establish the impact of a particular variable X on another variable Y in a system of at least three variables. The underlying question is: When is an association between two variables possibly causal and when is it the result of confounding? Wunsch (2007) points out that confounding is not a new issue.

In fact, in philosophy of science, Hans Reichenbach (1956) already discussed probabilistic causal structures where, for instance, two effects have a common cause and are therefore conditionally independent given the common cause. This led, more generally, to his 'common cause principle': if two events are correlated, either one causes the other, or they are both effects of a common cause—in Reichebach's words: "If improbable coincidence has occurred, there must exist a common cause" (Reichenbach 1956, p.157). In common cause structures, we say that Z "screens off" X from Y, as in figure 1.

At around the same time, statisticians were also aware that a correlation between two variables could be due to a common cause. Consider the correlation between smoking and lung cancer. One could suppose that one's (unknown) genotype could influence both smoking behaviour and the susceptibility to lung cancer. This explanation was proposed by R.A. Fisher when he was scientific consultant to the Tobacco Manufacturers' Standing Committee in the late 1950s (Fisher 1957), in order to counter the claim that smoking causes lung cancer. Thus, "without any direct causation being involved, both characteristics might be largely influenced by a common cause, in this case the individual genotype" (Fisher 1958).

Actually, an association between two variables X and Y can indeed be due to a common cause Z but it could also result from a causal relation between X and Y, *i.e.* X could be a cause of Y, or *vice-versa* Y could be a cause of X (as expressed by the principle of common cause). The distinction between a correlation due to a common cause Z or a correlation due to a causal relation between X and Y leads to the following definition of a confounder. According to Bollen (1989), Pearl (2000), and others, a confounding variable or confounder is a variable that is a common cause Z of both the putative cause X and its outcome Y. In this sense, confounding requires a suitably defined causal ordering of the variables. This definition leaves out the possibility of taking an intervening (intermediate or mediating) variable between the putative cause and the outcome as a confounder, though it is associated with the putative cause (as the latter has a causal influence on the former)

and it has an impact on the outcome.

Consider, in particular, the following two models corresponding to the directed acyclic graphs (DAGs) of Figure 1 (common cause variable) and Figure 2 (intervening variable). We recall that a DAG is an ordered set of nodes (variables) connected by edges (arrows) expressing a putative causal impact of the node at the head of the arrow on the node at the end of the arrow, and devoid of directed cycles such as $X \to Y$ and $Y \to X$ (Pearl, 2000). The directed nature of the DAG allow us to introduce time in the functioning of a causal structure. The variables appearing in the DAG need not be all simultaneous but may include, for instance, lagged variables.

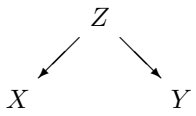

Figure 1: *Model 1: Common cause variable (confounder)*

Figure 2: *Model 2: Intervening variable (mediator)*

In Model 1, Z is a common cause of X and Y and therefore confounds the effect of X on Y, while Model 2 corresponds to a situation where Z is an intervening variable between X and Y.

$$Model\ 1\ [common\ cause\ variable]: \quad p_{X,Y,Z} = p_Z\, p_{X|Z}\, p_{Y|Z} \tag{1}$$

$$Model\ 2\ [intervening\ variable]: \quad p_{X,Y,Z} = p_X\, p_{Z|X}\, p_{Y|Z} \tag{2}$$

In both cases, although the two causal structures are quite different from one another, we have: $Y \perp\!\!\!\perp X \mid Z$. This shows that conditional independencies *alone* are insufficient for specifying the causal model. The latter requires background information on the putative causal network.

In the framework of directed acyclic graphs, Pearl (2000) has proposed a Back-Door Criterion in order to determine the set Z of confounders to be controlled for. Controlling, in this context, means keeping Z constant, or fixed, while examining the joint-variation between X and Y. In Pearl's view, a set of variables Z satisfies the Back-Door Criterion relative to the ordered pair of variables X and Y if:

(i) no node in Z is a descendant of X; and

(ii) Z blocks every path between X and Y that contains an arrow into X.

where blocking means — in the absence of a causal relation subsisting between the two — that two marginally dependent variables such as X and Y become independent once the modeller conditions on the third variable (or set of variables) Z, as in Models 1 and 2 above.

Considering condition (ii), the variable (or set) Z can be a common cause of X and Y (as in Model 1) or an intervening variable between X and Y (as in Model 2). However, the backdoor-criterion excludes controlling for an intervening variable between X and Y, as this variable would be a descendant of X (condition (i)). Condition (i) also excludes controlling for *colliders*, *i.e.* a common effect of two causes, as conditioning on a collider produces on the contrary a conditional dependency between the common causes (see also Hernán et al. 2002).

Historically, the concept of control is typically associated with experimental practices where manipulations are performed, for instance lab experiments or randomised studies. Here, controlling for a variable amounts to fixing the value of that variable such as fixing the level of atmospheric pressure in an experimental setting. In social science contexts, however, fixing the values of possible confounders is typically not feasible or ethical. Yet standard textbooks in the methodology of the social sciences, such as E. Babbie (2010), R.A. Jones (2000), C. Frankfort-Nachmias and D. Nachmias (2007), or in epidemiology, such as K.J. Rothman and S. Greenland (1998), insist on the need for controlling for possible confounders Z, when examining the impact of a variable X on a variable Y. This, it is usually argued, ensures that a change in Y is due to the change in X and not to the action of Z. Generally, standard textbooks recommend controlling for *all* variables possibly having an effect on Y, both confounders and covariates. However, the *methods* for controlling will usually be other than fixing and intervening.

To see how these recommendations are put to practical use in demography, we have examined as an example all articles that appeared in the *European Journal of Population*, one of the major journals in population science, during the past twelve months at time of writing, *i.e.* the four issues of 2012.

Broadly speaking, in most cases the authors follow the standard recommendations by including, based on a literature review, a vector of all known and observable determinants X of the effect variable Y into a single-equation model, either an event or duration model according to the data, and then considering the impact of each of these variables on Y independently from the other predictor variables, sometimes using a sequential approach taking into account each variable or group of variables at a time. A distinction is often made between key explanatory variables and controls (though not usually specifying if the latter are confounders, mediators, moderators, or independent covariates), less often between individual and contextual variables, and a theoretical framework is frequently developed in order to present the main research questions and hypotheses, though a full explanatory mechanism is most often absent. No causal ordering of the predictor variables is generally attempted, in the sense of a structure or mechanism responsible for the outcome variable;

this is the object of section 4. All predictor variables are implicitly considered in this approach as if they were independent from one another (except for multi-item scales), *i.e.* as if different structures of association among them had no impact on the generation of the outcome variable. Some authors do however use some form of similarity analysis to examine possible groupings among variables, and interactions between variables are sometimes examined. There are of course exceptions to this common approach. For example, Impicciatore and Billari (2012) use a multiprocess model composed of simultaneous equations for distinguishing causal effects from selection effects. And many other exceptions would be found if the set of texts were enlarged and additional population journals considered.

It is worth noting that the literature does not draw a substantial distinction between controlling in experimental settings and in observational settings. For instance, Pearl, in commenting on the proof of the Back-Door Criterion, says:

> When Z blocks all back-door paths from X to Y, setting $(X = x)$ [i.e., Pearl's *do* operator] or conditioning on $X = x$ has the same effect on Y. [...] Which means that the observation $(X = x)$ cannot be distinguished from the intervention $F_X = do(x)$. (Pearl 2000, p.80)

In other words, the *concept* of control is the same in experimental and observational settings, and, once we have (properly) performed an intervention or a conditioning, the result of the control will be the same.

3 Confounding and control in a directed network

Traditionally, confounding has been considered a threat for causal assessment when two variables X and Y are involved, the problem being that there may be a third variable Z such that the correlation between X and Y vanishes or is modified by the presence of Z. Here, instead, we reason on a *ordered network* of variables, and we consider the simplest case, in the form of three variables (X, Y, Z) where we specifically wish to evaluate the impact of X on Y, taking into account the action of Z. In this section, we do not discuss how the ordered network has been produced, we just assume a 3-variable model as given. Later in section 4, we discuss instead how we come to produce an ordered network in a structural modelling approach, and we will suggest that, in such a framework, confounding becomes a moot issue.

For expository purposes, we analyze in depth the three-variable case defined by a system that has been completely ordered, as represented by the DAG in Figure 3, or equivalently by a joint distribution structured as follows:

$$p_U(u \mid \theta) \;=\; p_{Y|Z,X}(y \mid z, x, \theta_{Y|Z,X})\; p_{X|Z}(x \mid z, \theta_{X|Z})\; p_Z(z \mid \theta_Z) \tag{3}$$

In this equation, the parameter θ represents the characteristics of the joint distribution of the $U = (X, Y, Z)$ and θ_Z, $\theta_{X|Z}$, $\theta_{Y|X,Z}$ stand for the characteristics of the corresponding distributions.

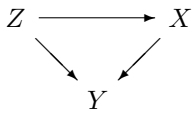

Figure 3: *3-component completely ordered system*

We show in particular that the status of Z as a confounder depends upon a complete specification of the model. We also give a formal presentation of the impact of confounding on the outcome factor. Our results provide an introduction to the complexities of higher-ordered systems.

The saturated case

Figure 3 and equation (3) represent the saturated case and suggest that Z causes X and (Z, X) cause Y. Here, Z is a confounding variable for the effect of X on Y. Z intervenes in two sub-mechanisms, those characterised by $p_{Y|Z,X}$ and $p_{X|Z}$. The effect of the variable X on the variable Y is described by analysing the variations of the distribution $p_{Y|X,Z}(y \mid x, z, \theta_{Y|X,Z})$, *i.e.* its characteristics $\theta_{Y|Z,X}$ (such as its conditional expectation and/or variance), relatively to different values of X, but this impact depends in general upon the values of Z.

Let us examine the consequences of neglecting the confounding variable Z in a causal analysis. This may be due either because of deficiency in background knowledge or because the confounder is not observed. It might then be tempting to collapse the diagram in Figure 3 into that of Figure 4.

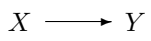

Figure 4: *2-component system*

Formally, Figure 4 may be obtained by integrating the latent, or neglected, variable Z out of (3):

$$p_{Y|X}(y \mid x, \theta_{Y|X}) = \frac{\int p_{Y|X,Z}(y \mid x, z, \theta_{Y|X,Z})\, p_{X|Z}(x \mid z, \theta_{X|Z})\, p_Z(z \mid \theta_Z)\, dz}{\int p_{X|Z}(x \mid z, \theta_{X|Z})\, p_Z(z \mid \theta_Z)\, dz} \quad (4)$$

$$p_X(x \mid \theta_X) = \int p_{X|Z}(x \mid z, \theta_{X|Z})\, p_Z(z \mid \theta_Z)\, dz \quad (5)$$

and therefore:

$$\theta_{Y|X} = f_1(\theta_{Y|X,Z}, \theta_{X|Z}, \theta_Z) \qquad \theta_X = f_2(\theta_{X|Z}, \theta_Z) \quad (6)$$

This boils down to evaluating the effect of X on Y through the conditional distribution $p_{Y|X}$. Here, the *prima facie* effect of X on Y is represented by $\theta_{Y|X}$ and is actually a complex combination of

three mechanisms involving $\theta_{Y|X,Z}, \theta_{X|Z}$ and θ_Z. Therefore, Figure 4 is an inadequate simplification of Figure 3. For instance, for policy purposes it would be misleading to base an intervention on differentiating, with respect to X, the conditional expectation of $(Y \mid X)$.

In the present case, the effect of X or of Z is evaluated through the parameters of the conditional distribution $(Y|X,Z)$. If interest focuses on the effect of X on Y, Z is a *confounding variable* that has to be controlled for .

We now examine situations where Figure 3, or equation (3), has been simplified by introducing restrictions represented by deleting one of the arrows, assuming some form of stochastic independence.

A first unsaturated case: $Y \perp\!\!\!\perp X \mid Z$

In this case, equation(3) simplifies into:

$$p_{X,Z,Y} = p_Z \; p_{X|Z} \; p_{Y|Z} \tag{7}$$

and Figure 3 becomes Figure 5.

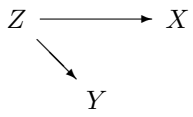

Figure 5: *A first unsaturated case*

Without any further assumption, X and Y would *not* be independent. Indeed, equation (4) becomes:

$$p_{Y|X}(y \mid x, \theta_{Y|X}) \quad = \quad \frac{\int p_{Y|Z}(y \mid z, \theta_{Y|Z}) \; p_{X|Z}(x \mid z, \theta_{X|Z}) \; p_Z(z \mid \theta_Z) \; dz}{\int p_{X|Z}(x \mid z, \theta_{X|Z}) \; p_Z(z \mid \theta_Z) \; dz} \tag{8}$$

reflecting that the association between X and Y is grounded on two sub-mechanisms represented by $p_{X|Z}$ and $p_{Y|Z}$. Nevertheless, this association between X and Y disappears once one conditions on the common cause and confounder Z, given that here $Y \perp\!\!\!\perp X \mid Z$.

In this case, the effect of X or of Z is evaluated through the conditional distribution of $Y|Z$, the latter being a simplification of $p_{Y|X,Z}$ through the supplementary condition $Y \perp\!\!\!\perp X|Z$. The causal effect of X on Y is null although X and Y are not marginally independent.

A second unsaturated case: $Y \perp\!\!\!\perp Z \mid X$

Figure 3 becomes Figure 6. Then variable Z is not a confounder anymore and $p_{Y|X}$ is obtained

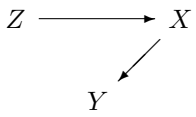

Figure 6: *A second unsaturated case*

directly from the decomposition of the joint distribution of $U = (X, Z, Y)$, equation (3), that factorizes into:

$$p_{X,Z,Y} = p_Z \; p_{X|Z} \; p_{Y|X} \tag{9}$$

In this case, the information on X is sufficient for predicting Y: adding information on Z would not improve the prediction on Y. Here, the effect of X on $Y$is correctly described by the characteristics embodied in the parameter $\theta_{Y|X}$ of the conditional distribution $p_{Y|X}$. Nevertheless, an intervention on Z would have an effect on Y, mediated by the value of X.

In this case, the effect of X (or of Z) is evaluated through the conditional distribution of $Y|X$, the latter being a simplification $Y|X, Z$ through the supplementary condition $Y \perp\!\!\!\perp Z|X$. The causal effect of X is measured through $\theta_{Y|X}$, whereas the value of Z should not be controlled for as its effect on Y is mediated by the value of X.

A third unsaturated case: $Z \perp\!\!\!\perp X$.

In this case, equation (3) simplifies into:

$$p_{X,Z,Y} = p_X \; p_Z \; p_{Y|X,Z} \tag{10}$$

and Figure 3 becomes Figure 7.

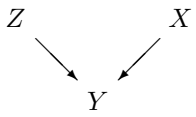

Figure 7: *A third unsaturated case*

The conditional distribution $p_{Y|X}$ becomes:

$$p_{Y|X}(y \mid x, \theta_{Y|X}) = \int p_{Y|Z,X}(y \mid z, x, \theta_{Y|Z,X}) \; p_Z(z \mid \theta_Z) \; dz \qquad where: \tag{11}$$

$$\theta_{Y|X} = f(\theta_{Y|Z,X}, \theta_Z) \tag{12}$$

Once again, Z is not a confounder anymore for the relation between X and Y: in particular, a change in the value of Z will not affect the value of X. Nevertheless, the effect of X on Y is disturbed by

the impact of Z, because variations in Y depend both upon variations in X *and* upon variations in Z.

In this case, the effect of X or of Z is evaluated through the conditional distribution of $Y|Z, X$. The marginal independence between Z and X has no bearing on the effect of Z or X on Y.

To summarize, the variable Z should be controlled for in the saturated case, in the first and third unsaturated cases and should *not* be controlled for in the second unsaturated case where X mediates the impact of Z on Y and where there is no direct impact of Z on Y, as $Y \perp\!\!\!\perp Z \mid X$.

In the previous paragraphs, we have shown on the basis of a simple multivariate example what variables the modeller should control for. Let us now consider briefly *how* to control for these variables. In the relevant cases considered in this section, controlling means examining the impact of a variation of X on the distribution of Y, holding the different values of Z fixed. For instance, in the example given at the very beginning of this paper, the impact of income on health would be measured according to each level of education (the different values of Z). These measures can possibly vary from one level of education to another, displaying an interaction or moderator effect of education on the income-health relation.

In practice, controlling for confounders can be done either at the design stage or at the analysis stage. At the design stage, the usual methods are randomization, restriction (ex-ante stratification), or matching. At the analysis stage one can have recourse to *e.g.* (ex-post) stratification, statistical modelling, or to instrumental variables in case of latent confounders. The interested reader is referred to dos Santos Silvas (1999) among others.

4 Confounding and control from a structural modelling perspective

In section 3 the structure of the joint distribution of all variables of interest is considered as given. In this section we introduce the framework of structural modelling as a strategy for providing an ordering of variables and for addressing the issue of confounding. A structural model is, essentially, a statistical model that provides a stochastic representation of a data generating process, interpreted as a global mechanism. A structural model decomposes such global mechanism in such a way that each component represents a sub-mechanism congruent with background knowledge and endowed with verified properties of structural stability.

For example, one can propose a global mechanism explaining self-rated health in Baltic countries, which includes different sub-mechanisms specifying the paths from *e.g.* locus of control via alcohol consumption to the outcome self-rated health (Gaumé and Wunsch 2010). The goal is now to analyse the impact of structurality on causal attribution. In section 4.1 we recall the main features of structural modelling. In section 4.2 we consider in particular the issues of confounding and control

in a structural framework.

4.1 Structural models: origins and main features

There are two main sources of a structural modelling approach, the econometric perspective and the social science one. In econometrics, the members of the Cowles Commission, in particular Koopmans (1950a) and Hood and Koopmans (1953), developed a concept of structural model under the motto 'no measurement without theory' (Koopmans 1947). A basic idea was to deduce the statistical implications of a model derived from economic theory deemed to represent the actual behaviors of economic agents. The recursivity of a model is showed to facilitate statistical inference and Herman Wold (1949, 1954), among others, defended the idea that a recursive model is a type of ideal model as long as individual agents tend to decide on one variable at a time, considering the other relevant variables as predetermined. In this model, the ordered structure of explanatory variables is given *a priori* by economic theory.

Another origin of structural modelling is rooted in the path analytic methodology developed by Sewall Wright and in the subsequent causal models for non-experimental research—from the 1960s and 1970s—of Blalock, Duncan, and Boudon, among others. The purpose here is to express a correlation between putative causes and effects through a recursive decomposition. They represented these relations in branching sequential order by an arrow-diagram, a graphic representation taken up much later and expanded by Judea Pearl especially, in his directed acyclic graphs (DAG) approach to causality (Pearl, 2000). Applications to linear models are presented in Pearl(2013).

Our starting point is a set of data X, possibly in the form of a data matrix. We are interested in a statistical model in the form of a set of probability distributions

$$\mathcal{M} = \{P_X^\theta \quad \theta \in \Theta\}$$

where θ is a parameter characterising a probability distribution and $\mathcal{M}$ represents a set of plausible hypotheses concerning the data generating process (DGP). We want the model to be "structural", meaning that the model should specify a plausible structure of the underlying DGP, relatively to a well-specified population of reference. Mouchart, Russo and Wunsch (2010) identify three main features of structural models:

(i) A recursive decomposition of the joint distribution interpretable as a sequence of sub-mechanisms, reflecting the causal ordering of the variables. Causal ordering is usually based on background knowledge including temporal information about the ordering of variables.

(ii) Congruence with background information.

(iii) Invariance or stability of the recursive decomposition across changes of the environment.

More specifically, once the vector of variables X is decomposed into an ordered sequence of p components, namely $X = (X_1, X_2, \cdots X_p)$ (with p typically much larger than 2), a recursive decomposition is a systematic marginal-conditional decomposition of the joint distribution of X, namely:

$$\begin{aligned} p_X(x \mid \theta) \quad = \quad & p_{X_p|X_1,X_2,\cdots X_{p-1}}(x_p \mid x_1, x_2, \cdots x_{p-1}, \theta_{p|1,\cdots p-1}) \\ & \cdot p_{X_{p-1}|X_1,X_2,\cdots X_{p-2}}(x_{p-1} \mid x_1, x_2, \cdots x_{p-2}, \theta_{p-1|1,\cdots p-2}) \cdots \\ & \cdot p_{X_j|X_1,X_2,\cdots X_{j-1}}(x_j \mid x_1, x_2, \cdots x_{j-1}, \theta_{j|1,\cdots j-1}) \cdots p_{X_1}(x_1 \mid \theta_1) \end{aligned} \tag{13}$$

where each $\theta_{j|1,\cdots j-1}$ stands for the parameter characterizing the corresponding conditional distribution $p_{X_j|X_1,X_2,\cdots X_{j-1}}$. The whole recursive decomposition is interpreted as the global mechanism in which several sub-mechanisms act. Thus the goal of a recursive decomposition is to endow the distributions P_X^θ with the interpretation that each component of the decomposition stands for one of the sub-mechanisms that compose the joint DGP of X. The recursive decomposition is built in such a way that the identified sub-mechanisms are interpretable from background knowledge; thus the order of the decomposition of X is crucial for the interpretability of the components as sub-mechanisms. Finally, invariance or stability of the model is required, as a major aim of structural modelling is to distinguish incidental from structural components of a data generating process.

What do we mean by 'mechanism'? We adopt the definition discussed by Illari and Williamson (2012): "A mechanism for a phenomenon consists of entities and activities organized in such a way that they are responsible for the phenomenon". This definition is general enough to be applicable to social contexts too. More specifically, it is compatible with an interpretation of the recursive decomposition in terms of mechanisms in a structural model (see for instance Wunsch, Mouchart, Russo 2012).

In this context, decomposing a global mechanism into a sequence of (autonomous) sub-mechanisms is tantamount to disentangling the action of each component in a sequence of sub-mechanisms operating in a global mechanism. We need to model *systems* of relations (rather than simple one-cause–one-effect relations) because in social science the typical case is multiple causes-multiple effects- complex systems. Thus a recursive decomposition carries explanatory power insofar as it decomposes a global mechanism into sub-mechanisms in the above sense. In this framework, the conditioning variables are interpreted as causes within a particular sub-mechanism, in which they have a specific function (see Mouchart and Russo (2011) and Wunsch, Mouchart and Russo (2012)).

4.2 Confounding and control in a structural framework

The issues of confounding and control are raised here in the context of defining and measuring the effect of an "ancestor" variable (*i.e.* ancestor-parent or ancestor-non parent), considered as a causing variable, on an outcome variable when the global mechanism is complex, *i.e.* involving

more than 3 variables. DAG terminology is used hereafter. From a structural modelling point of view, and its accompanying recursive decomposition, these issues have two facets. On the one side, the causing variable may be a parent in the last sub-mechanism of interest generating the outcome variable conditionally on its parents (the other ancestor variables being independent of the outcome conditionally on the parents) or may be an ancestor without being a parent, upstream in the causal chain. On the other side, when considering the sub-mechanism where a causing variable of interest is a parent, the effect of that variable also depends upon the level of the other parents in the case of interaction. Put otherwise, the issues of control and confounding are different when evaluating the effect of a non-parent ancestor on an outcome or when analyzing the effect of a parent, under the possibility of an interaction of effects of other parents.

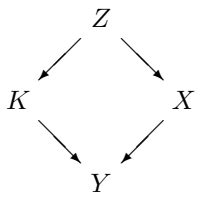

Figure 8: *A four-variate case*

Consider the Figures 3 and 8. In both cases, Z is a confounding variable for the effect of X on$Y$. We first want to control for Z when assessing the effect of X on Y and secondly, in a structural modeling approach, the analysis should not be restricted to the sole causal relation between X and Y: one should also examine the effect of Z on Y.

In the first case of Figure 3, the last sub-mechanism of interest is represented by the conditional distribution $p_{Y|X,Z}$. The *direct* effect of a variation of the variable X, from the value x to, say, the value $x + \delta$, is to cause a variation in the conditional distribution $p_{Y|X,Z}$ and "controlling for Z" means considering, in the conditional distribution $p_{Y|X,Z}$, a set of fixed values for Z in a neighborhood of a reference value of X.

Two approaches may be considered for measuring the direct effect of parent X on outcome Y in Figure 3. A first one evaluates explicitly a difference between the conditional distributions associated with the values of X, namely x and $x + \delta$. Thus if we write $d(k, q)$ for a discrepancy, *i.e.* a distance or divergence, between the distributions k and q, the effect of a variable X on Y controlling for Z may be measured through $d(p_{Y|X=x+\delta,Z=z},\ p_{Y|X=x,Z=z})$. In general, this is a function of x, z, and δ . Controlling for Z, for the effect of the variation of X from the value x to the value $x + \delta$, may be made by considering a sequence of possible values of Z, say $\{z_1, z_2, \cdots, z_r\}$, and interpreting the corresponding values of $d(p_{Y|X=x+\delta,Z=z_i},\ p_{Y|X=x,Z=z_i})$.

In a second approach, one may consider that the quantity $d(p_{Y|X=x+\delta,Z=z},\ p_{Y|X=x,Z=z})$ does not

capture the direct effect of interest, and the effect on some characteristics of the distribution $p_{Y|X,Z}$ is preferred. In other words, the direct effect of the variable X is evaluated through the parameter $\theta_{Y|X,Z}$ of the conditional distribution $p_{Y|X,Z}$. Thus, the effect on the conditional expectation may be evaluated by differentiating the conditional expectation $\frac{\partial E[Y|X=x,Z=z]}{\partial x}$. As an example, suppose:

$$E[Y \mid X = x, Z = z] = \alpha_0 + \alpha_1 x + \alpha_2 z + \alpha_3 xz$$

where α_3 measures an interaction of effects. Then

$$\frac{\partial E[Y \mid X = x, Z = z]}{\partial x} = \alpha_1 + \alpha_3 z$$

showing that this is a case where the effect of X depends upon the value of Z unless $\alpha_3 = 0$. Less trivially, if the conditional expectation is not linear but additive, *i.e.* :

$$E[Y \mid X = x, Z = z] = f_1(x) + f_2(z)$$

then $\frac{\partial E[Y|X=x,Z=z]}{\partial x} = f_1'(x)$ does not depend on z and is a function of x rather than a constant parameter.

Notice that this analysis of the direct effect is independent of the process generating the parents of the sub-mechanism of interest, *i.e.* independent of the dependence or independence or of a possible causal relation between X and Z. Thus the direct effect of X in Figures 5 and 7 is analyzed exactly in the same way whereas the case of Figure 6 is different because there Z is not anymore a confounder and is not a parent in $p_{Y|X,Z}$.

When analyzing the *total* effect of Z in a structural modeling approach, one observes that there are two paths from Z to Y: one direct and one indirect through X. This is precisely the concept of confounding in the sense of a common cause of X and of Y. The *direct* effect of Z on Y may be analyzed from $p_{Y|X,Z}$ exactly as the direct effect of X and operated under control of X: the confounding aspect of Z has no role for the analysis of the direct effect. The total effect of Z on Y takes into account that a variation of the value of Z from z to $z + \delta$ implies both a variation of the distribution $p_{Y|X,Z}$ *and* a variation in the distribution $p_{X|Z}$; this last one produces the *indirect* effect of Z on Y through X. The indirect effect is often measured, in the literature, by means of the difference between the total effect and the direct effect; in other words the total effect is equal to the direct effect in the absence of an indirect effect. In the case of the indirect path where X is a mediator between Z and Y, X should not be controlled for, contrary to the case of the direct path from Z to Y. For the effect of X on Y, the total effect is identical with the direct effect because there is only one path from X to Y.

Let us now consider Figure 8. Here, Z is at the source of two paths to Y, one through K and another one through X. Thus Z confounds two causal effects on Y: K on Y and X on Y. Figure 8 corresponds to the following recursive decomposition:

$$p_{Z,K,X,Y} = p_Z \; p_{K|Z} \; p_{X|Z} \; p_{Y|K,X}$$

which implies the following independence assumptions:

$$K \perp\!\!\!\perp X \mid Z \qquad\qquad Y \perp\!\!\!\perp Z \mid K, X$$

Here the sub-mechanism of interest is represented by $p_{Y|K,X}$; the direct effect on Y of X(or of K) may be analyzed exactly as in Figure 3, *i.e.* measuring the effect of X on Y for different values of K. Now, Z is not a parent anymore in $p_{Y|K,X}$ but is an ancestor non-parent of Y.
Suppose that, for a new observation, K is not observable but is known to have been generated by $p_{K|Z}$. Let us consider the effect on Y of X taking into account that Z is a common cause of K and of X and that K and X are both parents of Y. For this purpose, one may evaluate the conditional distribution $p_{Y|X,Z}$ as follows:

$$p_{Y|X,Z} = \int p_{K|Z}\ p_{Y|X,K}\ dK \tag{14}$$

This distribution does not represent a (structural) sub-mechanism but may be viewed as an approximation of $p_{Y|X,K}$ obtained when replacing K by the variable Z considered as a proxy for the sub-mechanism represented by $p_{K|Z}$. This (artificial) construction allows one to measure the effect on Y of a variation of X while controlling for Z, for instance by evaluating:

$$d(p_{Y|X=x,Z=z}, p_{Y|X=x+\delta,Z=z}) \qquad z \in \{z_1, z_2, \cdots\} \tag{15}$$

In this case, equation (15) represents an approximation of the effect of X on Y if the control of Z were a substitute for the control of K.

Suppose now that an intervention allows one to fix the value of X, rather than letting X be generated by $p_{X|Z}$, and that we want to measure the effect of a variation of Z while controlling for X. We might then evaluate, by means of relation (16), the effect of a variation of Z taking into account that K is randomly generated $p_{K|Z}$, although $p_{Y|X,Z}$ in equation (16) does not represent a sub-mechanism in the DGP:

$$d(p_{Y|X=x,Z=z}, p_{Y|X=x,Z=z+\delta}) \qquad x \in \{x_1, x_2, \cdots\} \tag{16}$$

Finally, we consider the total effect on Y of Z where a variation of Z will modify the conditional distributions of the parents of Y. Thus the total effect of Z on Y may be assessed through the effect of Z on $p_{K|Z}$ and on $p_{X|Z}$. But a "mean" effect may be evaluated through $p_{Y|Z}$ where:

$$p_{Y|Z} = \int\int p_{Y|K,X}\ p_{K|Z}\ p_{X|Z}\ dX\ dK \tag{17}$$

It should be stressed that $\theta_{Y|Z}$ is, in general, a complicated function of $\theta_{Y|K,X}$, $\theta_{K|Z}$ and $\theta_{X|Z}$ and that $p_{Y|Z}$ does not represent in itself a structural sub-mechanism but only a tool for assessing an ancestor effect.

In this section, we have shown, by way of simple examples, that the structural modeling approach, based on a recursive decomposition in terms of sub-mechanisms, provides an opportunity

for answering a series of relevant questions regarding different facets of cause-effect relations. In the absence of a structural approach, we believe that any answer to such questions is likely to be highly speculative.

5 Concluding Remarks

The problem of confounding and control arises when defining and measuring the effect of a cause on an outcome in the case of a complex system, requiring to distinguish between the total, direct and indirect effects. A confounder is traditionally defined as a common cause of both an effect and its possible determinant, and should be controlled for. This paper approaches the problem of confounding and control in the framework of structural modelling. In this framework, a complex, and therefore multivariate, global mechanism is recursively decomposed into an ordered sequence of sub-mechanisms represented by conditional distributions under conditions. The recursive decomposition allows the modeller to interpret the conditioning variables as causes of the outcome. In this framework, the effect of a variation of a cause is to produce a variation of the conditional distribution of the outcome, and the problem of control consists in analyzing this co-variation for different values of the variable to be controlled for.

This paper presents a general framework for the analysis of causal attribution. The role of time is not treated explicitly, although time is an essential component in causal analysis. In a structural modelling framework, the recursive decomposition provides space for making explicit sub-mechanisms where time plays a role in the ordering of variables.

For an outcome — and accordingly a sub-mechanism — of interest, the recursive nature of the decomposition leads to distinguishing between parents and ancestors of the outcome and to recognize the possibility of multiple paths between a variable appearing in the global mechanism and the outcome of interest. In particular, confounders can be seen as variables that are at the origin of two or more different paths leading to the outcome (as in figure 8) and may accordingly have both a direct and an indirect effect on an outcome. Confounders do not appear *per se* in the system but are subsumed within the larger framework of interrelations among variables in the structural model. In a structural perspective, confounding is therefore a moot issue because a structural model should incorporate the multiple paths (*e.g.* direct and indirect) leading from the causes to the outcome in the various sub-mechanisms, thus taking possible confounders into account.

From the point of view of DAGs, we have to include in the causal analysis all the variables having an arrow leading to the outcome variable, and their relevant ancestors, as suggested by background knowledge. When analyzing the effect of a causing variable on the basis of the sub-mechanism generating the outcome of interest, one should condition the outcome on all its parents, and not just on the confounders, as detected by Pearl's Back-Door Criterion. This goes beyond the

typical approach, that takes into account only those variables with an arrow into the cause that are confounders by application of Pearl's Back-Door Criterion.

Following Russo (2009), the effect of a cause is to modify the conditional distribution representing the sub-mechanism generating the outcome of interest. In other words, a variation of the cause produces a variation of the conditional distribution of the outcome variable. Thus the measure of the effect is consistently conveyed in terms of the *discrepancy* (*i.e.* distance or divergence) between the conditional distributions relative to different values of the causing variables.

Let us write $d(k, q)$ for a discrepancy, *i.e.* a distance or divergence, between the distributions k and q. Thus, in Figure 3, the direct effect of a variable X on Y controlling for Z may be measured through $d(p_{Y|X=x+\delta,Z=z}, p_{Y|X=x,Z=z})$. In general, this is a function of x, z, and δ. The choice of the discrepancy function is arbitrary, though different measures operationalize a same concept of effect.

Often, the quantity $d(p_{Y|X=x+\delta,Z=z},\ p_{Y|X=x,Z=z})$ does not capture the direct effect of interest, and the effect on some characteristics of the distribution $p_{Y|X,Z}$ is preferred. Thus the relevant characteristic may be the conditional expectation, or the conditional variance as in ARCH models widely used in financial econometrics, or the tail of the conditional distribution as in many actuarial models. Simplifying assumptions are often introduced but sometimes not always explicitly. For instance, under a complete normality assumption, with linear and homoscedastic regression, the discrepancy is simply characterised by the coefficient of the (putative) cause in the regression of the outcome on its parents. But this is an extremely (and often abusively) simple case: the effect on the conditional expectation may be taken into account by a non linear regression function with non-additivity, and more generally, with interaction effects.

Thus controlling means examining the behaviour of conditional distributions representing sub-mechanisms, under different values of the control variable. These different values may however result from two different procedures. A first procedure, often used with observational data, conditions on different values of the control variable. Let us call it "controlling by conditioning". Another procedure intervenes in the global mechanism by fixing the values of the variable to be controlled for. Let us call it "controlling by intervention". It should be noted that controlling by intervention implies a modification of the DGP: the variable to be controlled for is not anymore generated by the sub-mechanism identified in the recursive decomposition, but by the intervention itself. Such modification of the sub-mechanism may also lead to modifying other sub-mechanisms, in particular when the intervention is operated under a change of policy active on the global mechanism.

There aren't any statistical test or conditional independencies that *alone* tell us whether a variable is a confounder or whether a confounder has to be included in the model. Such decisions are taken, to the best of our knowledge, on the basis of background information and of preliminary analyses of data. "To the best of our knowledge" implies keeping open the possibility of improving the structural model by incorporating innovations in the field of data, theory, methods. In particular,

the progress of knowledge can lead to uncovering previously unrecognised confounding variables and unrecognised mechanisms.

Acknowledgements. The authors thank Joniada Milla for valuable comments and suggestions.

References

Babbie E. R. (2010), *The practice of social research*, Wadsworth, Belmont.

Bollen K.A. (1989), *Structural Equations with Latent Variables*, John Wiley & Sons, New York.

Dos Santos Silva, I. (1999), *Cancer Epidemiology: Principles and Methods*, chapters 13 and 14, IARC, WHO, Lyon.

Fisher R.A. (1957), Alleged dangers of cigarette smoking, *British Medical Journal*, II, 297-298.

Fisher R.A. (1958), Lung cancer and cigarettes, *Nature*, 182, July 12, p. 108.

Frankfort - Nachmias C. and D. Nachmias (2007), *Research methods in the social sciences*, Worth Publishers, Gordonsville.

Gaumé C. and Wunsch G. (2010), Self-rated Health in the Baltic Countries, 1994-1999, *European Journal of Population*, **26**(4): 435-457.

Hernán M.A., Hernández-Diaz N., Werler M.M., and Mitchell A.A. (2002), Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology, *American Journal of Epidemiology*, **155**(2), 176-184.

Hood, W.C. and T.C. Koopmans (eds) (1953), *Studies in Econometric Methods*, Cowles Founfation Monograph 14, John Wiley & Sons, New-York.

Illari, P. M. and Williamson, J. (2012), What is a mechanism? Thinking about mechanisms across the sciences, *European Journal for Philosophy of Science*, 2 (1), 119-135.

Impicciatore, R. and F. Billari F., Secularization, Union Formation Practices, and Marital Stability: Evidence from Italy, *European Journal of Population*, **28**(2), 119-138.

Jones R. A. (2000), *Méthodes de recherche en sciences humaines*, De Boeck, Brussels.

Koopmans T.C. (1947), Measurement without Theory, *Review of Economic Statistics*, **29**, 161-173.

Koopmans T.C. (ed.) (1950), *Statistical Inference in Dynamic Economic Models*, Cowles Founfation Monograph 10, John Wiley, New-York.

Mouchart M. and Russo F. (2011), Causal explanation: recursive decompositions and mechanisms, in P. McKay Illari, F. Russo, and J. Williamson (eds), *Causality in the sciences*, Oxford University Press, Oxford, 317-337.

Mouchart M. , F. Russo and G. Wunsch (2010), Inferring Causal Relations by Modelling Structures, *Statistica*, **LXX**(4), 411-432.

Pearl J. (2000), *Causality*, Cambridge University Press, Cambridge.

Pearl J. (2013), Linear models: a useful "microscope" for causal analysis, *Journal of Causal Inference*, **1**(1), 155-170.

Reichenbach H. (1956), *The direction of Time*, University of California Press. Reprinted (2000), Dover Publications, Mineole, New York.

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

Russo F.(2009), *Causality and Causal Modelling in the Social Sciences: Measuring Variations*, Methodos Series Vol.5, Springer.

Wold, H.O. (1949), Statistical estimation of economic relationships, *Econometrica* **17** Supplement: Report of te Washington Meeting, 1-22.

Wold, H.O. (1954), Causality and Econometrics, *Econometrica* **22**(2), 162-177.

Wunsch G. (2007), Confounding and Control, *Demographic Research*, **16**, pp. 15-35.

Wunsch G., Mouchart M. and Russo F. (forthcoming), Functions and mechanisms in structural-modelling explanation, *Journal for General Philosophy of Science*.