Controlling Variables in Social Systems

A structural modelling approach

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

[a] *CORE and ISBA, Université catholique de Louvain*

[b] *Demography, Université catholique de Louvain*

[c] *Philosophy of Science, University of Amsterdam*

June 7, 2016

Abstract

Background. In demography, and in other social sciences, determining the variables to be controlled for is usually a major problem when analyzing possible causal relations. Clear guidelines in the literature are scarce and actual practice is often questionable.

Objectives. A structural modeling approach is presented as a consistent framework for determining a coherent set of guidelines for deciding what variables should be controlled.

Methods. The method is based on a recursive decomposition of the multivariate distribution, represented by a directed acyclic graph and reflecting the causal mechanism and sub-mechanisms of the data generating process.

Results. Two rules are developed for determining control variables when studying respectively the *direct* and the *total* effects of a cause on an outcome. The rules can easily be applied in the framework of a causal model based on background knowledge and invariant to changes of the environment.

Conclusions. Our approach for determining control variables is simpler and more consistent than the alternative ones based on Pearl's back-door criterion. It takes into account both confounders and other immediate causes leading to variations in the distribution of the outcome variable and possibly being in interaction with the causal variable

Keywords: Causality, Control, Causal Modelling, Structural Modelling, Recursive Decomposition, Total Effect, Direct Effect, Directed Acyclic Graph.

Corresponding Author: Michel Mouchart, ISBA, Université catholique de Louvain, 20 Voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium. *e-mail*: michel.mouchart@uclouvain.be

Contents

1 Introduction

As stated in the *Glossary of Statistical Terms* (2014), there are at least three senses of "control" in statistics: a member of a control group to whom no treatment is given, a controlled experiment, and to control for a possible confounding variable. According to A. Crossman (2014), a control variable is a variable whose value is held constant in a research analysis in order to answer four basic kinds of questions (brackets ours):

1. Is an observed relationship between two variables just a statistical accident? (*i.e.* is a correlation spurious?)

2. If one variable has a causal effect on another, is this effect a direct one or is it indirect with other variable intervening?

3. If several variables all have causal effects on the dependent variable, how does the strength of those effects vary? (and how do several causal variables have a possible combined effect on the outcome variable?)

4. Does a particular relationship between two variables look the same under various conditions?

Historically, the issue of control is typically associated with experimental practices where manipulations are performed, for instance in laboratory experiments or in randomized studies. R.A. Fisher (1935) can be considered the modern theorizer of a tradition that traces back to, at least, J.S. Mill (1843). In this context, controlling for a variable amounts to holding constant the value of that variable, *e.g.* repeating an experiment under different fixed levels of atmospheric pressure. In non-experimental or observational sciences, such as demography, where controlled experiments are often impossible or unethical, 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), generally recommend controlling for the variables possibly having an effect on an outcome variable Y, either by intervening, when feasible, or by conditioning in a statistical model. In a non-experimental context, holding constant the value of a control variable may be operationalized in two different ways. Either, the data are split according to different values of the control variable and the characteristics of the different subsamples are examined, or a model is first specified and the properties of the distributions are examined under different values of the control variable.

To see how the above recommendations are put into practice in a non-experimental science, a perusal of some recent articles in demography (see Wunsch, Mouchart, Russo, 2014) shows that, broadly speaking, in most cases the authors follow the standard recommendations. They include, based on a literature review, a vector of all known and observable possible determinants X of the outcome variable Y into a single-equation model and then consider the impact of each of these

variables on the outcome Y, the other predictor variables being fixed. A distinction is often made between key explanatory variables and other variables to be controlled for. It is however not often specified whether those other variables are confounders, mediators, moderators, or independent covariates, neither is a distinction usually made between individual characteristics and contextual variables. Moreover, 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.

In this sample of articles, no causal ordering of the predictor variables is generally attempted, in the sense of a structure or mechanism responsible for the outcome variable. 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 and many other exceptions would be found if the set of texts were enlarged.

The situation seems rather similar in other fields. For example, in organizational research, T. E. Becker (2005) has shown that among a random sample of articles he examined, published between 2000 and 2002, more than half the articles provided no explanation for including one or more control variables. Even when justification is provided, often no explanation of the nature of the relationship between control and outcome variables in offered (Atinc, Simmering and Kroll, 2012). Actually, following Spector and Brannick (2011), many papers examined used control variables inappropriately. The present paper focuses on the issue of determining which variables should be controlled for when identifying the direct or the total effect of a cause in a complex system.

The previous selection of the literature follows to a large extent what H.-P. Blossfeld (2009) has called the *causation as robust dependence* approach. In this view, as discussed by D. R. Cox (1992), a variable X is a plausible cause of another variable Y if the dependence between the two cannot be eliminated by introducing additional variables in the analysis. The problem with this approach is that it is impossible to be sure that all relevant variables have been controlled for. Moreover, as H.-P. Blossfeld stresses, because covariates are often correlated, parameter estimates depend upon the specific set of variables included in the statistical model.

In order to go beyond this causation *via* association approach, as D.R. Cox (1992) has called it, and to be able to answer among others the questions raised by A. Crossman at the beginning of this section, one needs what H.-P. Blossfeld (2009) has coined a *causation as generative process* approach. In other words, one should characterize the properties of the underlying data generating process, *i.e.* the mechanism behind the data. More generally, in a perspective of explanation and policy intervention, as opposed to a purely descriptive point of view, one requires understanding the plausible mechanism and sub-mechanisms generating the data in a particular context and during

a specific period of time. This paper, at variance from a concept of causality taken as a primitive concept represented by a directed graph, makes use of a concept of causality derived from a specific structural modelling approach and based on the identification of mechanisms and sub-mechanisms.

Russo (2009) emphasizes that assessing the effect of a causal variable on an outcome variable amounts to assessing the *variation* that the cause-variable brings about in the conditional distribution of the outcome variable (admittedly, variations in the cause-variables can be observed or 'provoked' by manipulation—see Russo (2014)). More specifically, at a somewhat abstract level, the effect should be measured by the discrepancy, *i.e.* distance or divergence[1], between the conditional distributions of the outcome variable relative to different values of the causal variable. Typically, modellers measure the discrepancy between the same characteristics, or parameters, of the two conditional distributions. It will be shown that in complex social systems, controlling is an unescapable issue when measuring the effect of a cause.

As the justification for inclusion of control variables is very often limited or absent in many studies (Schjoedt and Bird, 2014), the purpose of this paper is to examine the issue of control in the perspective of a specific structural modelling of complex social systems of variables and relations, where there are multiple causes and multiple effects. In particular, this paper shows what are the relevant variables that have to be controlled for when studying cause-effect relations in social systems. The present paper does not deal however with methods of data collection and analysis.

The order of exposition is as follows. Section 2 considers the issue of control in a structural modelling perspective based on a recursive decomposition of a joint distribution standing for an explicit mechanism or data generating process and represented by a Directed Acyclic Graph (DAG). Section 3 deals with the simplest case of three variables in a directed network and examines in this framework the variables that have to be controlled for. We then consider the issue of control in more complex systems (Section 4). We derive two simple rules for selecting the variables to be controlled for in the cases of direct and total effects, without having recourse to Pearl's back-door criterion (see *e.g.* Pearl, 2000). We conclude (Section 5) that the real issue at stake is the modelling strategy, rather than the search for robust statistical associations. The paper is complemented by an Appendix discussing the non-uniqueness relations between DAG's and recursive decompositions.

2 Controlling from a structural modelling perspective

In this section, the framework of structural modelling is presented as a strategy for providing an ordering of variables underlying the recursive decomposition required for attributing causes and effects and for determining the control variables. In this perspective, a structural model is a statistical

[1]Roughly speaking, a divergence between distributions shares most of the properties of a distance but does not satisfy neither the symmetry nor the triangular inequality one.

model that provides a stochastic representation of a Data Generating Process (DGP) interpreted as a global mechanism. In order to understand the functioning of this global mechanism, the latter is decomposed into an ordered sequence of sub-mechanisms congruent with background knowledge and endowed with properties of invariance or structural stability (for a discussion, see also Russo, 2014). Causal attribution is then based on such a structural model. Section 2.1 recalls some origins of structural modelling and Section 2.2 presents its main features.

The structural model, as presented in this paper, is substantially different from that developed in the literatures on Structural Equation Models (SEM) in econometrics and in social science. In particular, the approach of this paper is not based on a system of equations, but rather on an analysis of a multivariate distribution. Moreover, the modelling stage is essentially distribution-free and the hypotheses of conditional independence are σ-algebraic in nature. The distributional hypotheses are accordingly introduced only at the estimation stage. More substantially, the usual SEM approach endeavors at *representing* causal knowledge while the structural perspective presented here aims at *constructing* causal knowledge. Thus the emphasis, in this paper, on mechanisms and sub-mechanisms, on background knowledge and on structural stability.

2.1 Structural models: origins

The specific approach to structural modeling, as presented in the next section, is grounded on two main streams, in econometrics on the one hand and in social science on the other hand. In econometrics, the members of the Cowles Commission, in particular Koopmans (1950) 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 was shown to facilitate statistical inference. Herman Wold (1949, 1954), among others, considered 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 approach, the ordered structure of explanatory variables is given *a priori* by economic theory. In this view, simultaneity of equations was viewed as an effect of incomplete observability, in particular of time aggregation. Recently, the structural approach in econometrics has known a substantial revival of interest, in particular within a debate concerning a so-called "a-theoretical econometrics" (see *e.g.* Fagiolo, Moneta and Windrum, 2007).

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 in the social sciences—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 later and expanded by Judea Pearl with his Directed Acyclic Graphs (DAG) approach to causality (see *e.g.* Pearl, 2000).

2.2 Structural models: a specific approach

The starting point of a structural model is a set of variables X. We consider a statistical model in the form of a set of probability distributions

$$\mathcal{M} = \{P_X^\theta \quad \theta \in \Theta\} \tag{1}$$

where θ is a parameter characterising a probability distribution and $\mathcal{M}$ represents a set of plausible hypotheses concerning the data generating process (DGP). Representing the DGP by probability distributions characterized by a parameter, implies that what is "explained" by the statistical model is embodied in the parameter whereas what is not explained is embodied in the stochastic component of the probability distributions.

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 a structural model:

(i) A recursive decomposition of the joint distribution interpretable as a sequence of sub-mechanisms, reflecting the causal ordering of the variables underlying the putative mechanism.

(ii) Congruence with background information: causal ordering of the variables is usually based on prior knowledge, including temporal information about the ordering of variables.

(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:

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

where each $\theta_{j|1,\cdots j-1}$ stands for the parameters characterizing the corresponding conditional distribution $p_{X_j|X_1,X_2,\cdots X_{j-1}}$. The whole recursive decomposition is interpreted as a global mechanism decomposed into an ordered sequence of acting sub-mechanisms. For the sub-mechanisms to act autonomously, one should also add a condition of mutual independence among the parameters

$\theta_{j|1,\cdots j-1}$, *i.e.* these parameters should be variation-free in a sampling theory approach or a priori independent in a bayesian approach; this topic has been considered in Mouchart, Russo and Wunsch (2010, 2011).

The recursive decomposition is the cornerstone of the explanatory power of a structural model because it endows the distribution P_X^θ with the interpretation that each component of the decomposition, *i.e.* the distribution of an outcome variable conditional on its explanatory variables, 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. 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 structural from incidental components of a data generating process.

In general, a recursive decomposition may be represented by a DAG (Pearl 1988). It should however be noted that the correspondence between DAGs and recursive decompositions is not one-to-one. The Appendix gives a simple example of a DAG corresponding to 2 different recursive decompositions that may be contextually different but correspond nevertheless to a same joint distribution of the variables and to a same specification of variables to be controlled for.

What do we mean by 'mechanism'? In the words of 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 sub-mechanisms in a structural model (see for instance Wunsch, Mouchart, Russo (2014)).

In this context, decomposing a global mechanism into a sequence of autonomous sub-mechanisms is tantamount to disentangling the function of each component in a sequence of sub-mechanisms operating in a global mechanism. We specifically 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 or role (see Mouchart and Russo, 2011; Wunsch, Mouchart and Russo, 2014).

Because a structural model, as proposed in this paper, is based on the concepts of mechanism, sub-mechanisms and functions, a characteristic feature is its ability to identify, in particular, which variables have a direct effect on each outcome considered and to deduce therefrom which variables should be controlled for, as we shall see in Sections 3 and 4. The structural approach allows one not only to evaluate the causes of an outcome but also the effects of a cause on several outcome variables. In Section 5, we will however discuss some shortcomings and limitations of this approach.

3 Controlling in the simplest case of 3 variables

For expository purposes, we analyze in depth the three-variable case (X, Y, Z) defined by a system that has been completely ordered on the basis of a structural model and represented by the *Directed Acyclic Graph* (DAG) of Figure 1, or equivalently by a joint distribution decomposed 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.

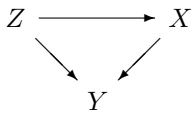

Figure 1: *3-variable completely ordered system*

In the present paper, the focus is on the analysis of the effects of causes on outcome variables in saturated and unsaturated models. This is achieved by examining the distributions of variables considered as outcomes conditionally on their *ancestors*, *i.e.* antecedent variables in the causal sequence. An immediate cause of an outcome is called a *parent* of this outcome, the latter being called a *child* of the cause. The analysis is therefore congruent with the order of the recursive decomposition.

In this structural modelling approach, the left-hand side of (3) represents an overall mechanism whereas the three terms on the right-hand side of (3) represent a decomposition of the overall mechanism into three sub-mechanisms.

The effect of a parent, for instance X or Z, on the outcome variable Y, is called a *direct* effect. This corresponds to the final sub-mechanism generating Y and is determined by the conditional distribution $p_{Y|Z,X}$, *i.e.* by $\theta_{Y|Z,X}$ representing its characteristics, such as its conditional expectation and/or variance. Said differently, the direct effect is evaluated from the decomposition $p_{X,Y,Z} = p_{X,Z}\ p_{Y|Z,X}$ independently of any hypothesis on $p_{X,Z}$. The effect of a parent, say X, on the outcome variable Y, evaluated from the conditional distribution $p_{Y|X}$, *i.e.* neglecting the other parents of the outcome, in this example Z, will be called a *prima facie direct* effect. Later on, the effect of a non-parent ancestor will be called a *total* effect. It can be due to multiple directed (causal) paths from the cause to the outcome of interest or to a sole indirect path. These effects will be detailed under different hypotheses.

In the present Section 3, we examine how different issues of controlling for X or for Z may arise when evaluating the effects on the outcome variable Y. We show in particular that the status of Z,

or of X, as a control variable, as defined in the Introduction, depends upon a complete specification of the model. This section provides an introduction to the complexities of higher-ordered systems.

3.1 The saturated case

Figure 1 and equation (3) represent a saturated case because, from a probabilistic point of view, equation (3) represents an identity without underlying restrictions, excepting the order of the variables. Heuristically, the DAG representing a saturated case is such that adding any new directed arrow would create a cycle in the graph. In this example, Z causes X and (Z, X) cause Y. In this case, X is a parent of Y whereas Z is both a parent and a non-parent ancestor of Y. Indeed, there is a direct path $Z \to Y$ and an indirect path $Z \to X \to Y$ from Z to Y.

3.1.1 The effect of X on Y

Graphically, there is only one direct path from X to Y and Z is a *confounding* variable for the effect of X on Y, as it is a common cause of both X and Y. For a discussion of the concept of confounding variable or confounder, see *e.g.* Bollen (1989) and Wunsch (2007). Variable X is a *mediator*, or intermediate variable, on the *indirect path* from Z to Y. The variable Z appears in two conditional distributions, or sub-mechanisms, of the right-hand side of (3), namely $p_{Y|Z,X}$ and $p_{X|Z}$.

The *direct effect* of the variable X on the variable Y can be described by analysing the variations of the distribution $p_{Y|X,Z}(y \mid x, z, \theta_{Y|X,Z})$ relatively to different values of X, but this impact depends in general upon the values of Z. For this reason, the confounding variable Z should be controlled for. This means that the variation of the distribution $p_{Y|X,Z}$ should be examined for different values of Z.

As the direct effect of X on Y is characterized by the conditional distribution $p_{Y|X,Z}$, it is important to condition the outcome on both X *and* Z in order to detect cases with *interaction*, where the direct effect of X depends upon the value of Z, as opposed to cases without interaction where the direct effect of X is not affected by the value of Z. An interaction effect is also called a *moderator* effect, especially in the psychological literature (Baron and Kenny, 1986). An interaction effect cannot be adequately represented in a DAG. An interaction may be due to an intrinsic non-additivity of direct effects or to neglecting, in a model, the action of other variables, as might occur when the action of these variables is unknown to the model builder or when these variables are not observable. Moreover, detecting an interaction may be a subtle issue because it may depend upon some analytical features of the model. For instance, if the conditional expectation of Y has the form $\exp\{\alpha_0 + \alpha_1 X + \alpha_2 Z\}$, differentiating the conditional expectation reveals an interaction effect whereas the log of the conditional expectation is simply additive without interaction.

The *prima facie direct effect* of X on Y, captured by the conditional distribution $p_{Y|X}$, may be viewed as an examination of the consequences of neglecting the confounding variable Z in a causal analysis. This may be due to deficiency in background knowledge or because the confounder is not observed. Formally, Figure 2 may be viewed as a result of collapsing the diagram of Figure 1 by

Figure 2: *2-variable system*

eliminating the variable Z and corresponds to integrating the latent, or neglected, variable Z out of (3):

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

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

and therefore:

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

Here, the *prima facie* direct effect of X on Y is represented by $\theta_{Y|X}$ and is actually a complex combination of three sub-mechanisms involving $\theta_{Y|X,Z}, \theta_{X|Z}$ and θ_Z. Therefore, Figure 2 is a misleading simplification of Figure 1. For instance, for policy purposes it would be wrong to base an intervention on differentiating, with respect to X, the conditional expectation of $(Y \mid X)$. Finally, the difference between the *prima facie* direct effect and the direct effect may be interpreted as a supplementary effect of the confounder Z.

As a conclusion, controlling for Z is required for evaluating the direct effect of X but not for the *prima facie* direct effect of X. However, the distribution $p_{Y|X}$ may be misleading for understanding the role of X in the sub-mechanism generating Y.

3.1.2 The effect of Z on Y

Graphically, there are two paths from Z to Y, a direct path and an indirect one through the variable X. In this case, the effect of Z on Y, captured by $p_{Y|Z}$, is usually called the *total* effect transmitted here by the direct and indirect paths, or more generally by multiple directed (causal) paths from cause to outcome. In this example, the total effect and the *prima facie* direct effect of Z on Y are identical.

We distinguish a *total effect* of Z on Y, captured by the conditional distribution $p_{Y|Z}$, and a *direct effect* captured by the conditional distribution $p_{Y|X,Z}$. Several remarks are in order. Firstly, the

conditional distribution $p_{Y|Z}$ does not represent a sub-mechanism but is a mixture of several sub-mechanisms, in this case two:

$$p_{Y|Z} = \int p_{Y|X,Z}\, p_{X|Z}\, dX \tag{7}$$

Secondly, as for the direct effect of X, the direct effect of Z on Y may be with or without *interaction*. Thus the variable X should be controlled for by examining the direct effect of Z for different values of X. Thirdly, the supplementary effect of Z on Y, generated by the indirect path $Z \to X \to Y$ and defined as the difference between the total effect and the direct effect, depends or not upon the value of X according to whether the direct effect of Z is with or without interaction with X. Take, for instance, the very simple case where effects are measured by means of conditional expectations in a gaussian situation where $\mathbb{E}(Y \mid X, Z) = \alpha_0 + \alpha_1 X + \alpha_2 Z + \alpha_3 XZ$ and $\mathbb{E}(Y \mid Z) = \beta_0 + \beta_1 Z$. In this case, the supplementary effect of Z is $\beta_1 - (\alpha_2 + \alpha_3 X)$ and therefore depends, or not, on X according to whether α_3 is different or equal to 0.

3.2 Unsaturated cases

We now examine situations where Figure 1, or equation (3), has been simplified by introducing restrictions represented by deleting one of the arrows, corresponding to some form of stochastic independence.

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

In this case, equation(3) simplifies into:

$$p_{X,Z,Y} = p_Z\, p_{X|Z}\, p_{Y|Z} \tag{8}$$

and Figure 1 becomes Figure 3.

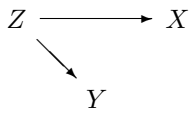

Figure 3: *A first unsaturated case*

The effect of X on Y

Without any further assumption, X and Y would *not* be independent. Indeed, the association between X and Y is captured by equation (4) transformed as follows:

$$p_{Y|X}(y \mid x, \theta_{Y|X}) \quad = \quad \frac{\int p_{Y|Z}(y \mid z, \theta_{Y|Z})\, p_{X|Z}(x \mid z, \theta_{X|Z})\, p_Z(z \mid \theta_Z)\, dz}{\int p_{X|Z}(x \mid z, \theta_{X|Z})\, p_Z(z \mid \theta_Z)\, dz} \tag{9}$$

reflecting that this association is grounded on the combined action of two sub-mechanisms represented by $p_{X|Z}$ and $p_{Y|Z}$ but disappears however once one conditions on the common cause and

confounder Z, given that here $Y \perp\!\!\!\perp X \mid Z$. The direct effect of X on Y is actually null although X and Y are not marginally independent.

The effect of Z on Y

Under the hypothesis $Y \perp\!\!\!\perp X \mid Z$, the direct effect of Z on Y is captured by the sub-mechanism $p_{Y|Z}$, independently of X. The variable X should therefore not be controlled for. If, however, one did control X by examining the joint distribution of Y and Z for different values of X, namely:

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

one would find a same distribution $p_{Y|Z}$ for each value of x_i but, in the case of a discrete variable X, one would reduce the number of observations in the sub-samples corresponding to each value x_i.

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

In this case, equation(3) simplifies into:

$$p_{X,Z,Y} = p_Z \; p_{X|Z} \; p_{Y|X} \tag{11}$$

and Figure 1 becomes Figure 4.

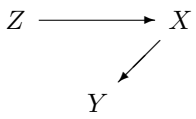

Figure 4: *A second unsaturated case*

The effect of X on Y

Variable Z is not a confounder anymore and $p_{Y|X}$ is obtained directly from the decomposition of the joint distribution given by (11). 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 direct effect of X on Y is correctly described by the characteristics of the conditional distribution $p_{Y|X}$. In this situation, the direct effect of X is evaluated through $\theta_{Y|X}$, and the value of Z should not be controlled for as its effect on Y is mediated by the value of X.

The effect of Z on Y

In this case, there is no direct effect of Z on Y because $p_{Y|X,Z} = p_{Y|X}$ and the effect of Z on Y is completely mediated by X. Moreover, for the total effect of Z, the variable X should not be controlled for because the distribution $p_{Y|Z}$ is actually a blending of two sub-mechanisms, namely

$$p_{Y|Z} = \int p_{Y|X} p_{X|Z} \, dX \tag{12}$$

where X is an active random variable.

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

In this case, equation (3) simplifies into:

$$p_{X,Z,Y} = p_X \; p_Z \; p_{Y|X,Z} \tag{13}$$

and Figure 1 becomes Figure 5.

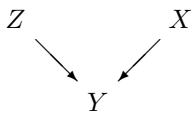

Figure 5: *A third unsaturated case*

The effect of X or Z on Y

In this case, the roles of X and of Z are perfectly symmetrical. The direct effect of X (or of Z) is evaluated through the conditional distribution of $Y|Z,X$ and the marginal independence between Z and X has no bearing on the direct effect of Z or X on Y. Changes in the distribution of Y are due, in this model, to changes in X and/or in Z. Here Z (or X) is not a confounder for the relation between X (or Z) and Y; however, variables X and Z can have an interaction effect on Y, meaning that the effect of X on Y depends upon the value of Z, and vice-versa. For this reason, it is necessary to control for Z (or X) when studying the impact of X (or Z) on the distribution of Y. As X and Z are independent, the *prima facie* direct effect of X, given by the conditional distribution $p_{Y|X}$, becomes:

$$p_{Y|X}(y \mid x, \theta_{Y|X}) = \int p_{Y|Z,X}(y \mid z, x, \theta_{Y|Z,X}) \; p_Z(z \mid \theta_Z) \; dz \qquad where: \tag{14}$$

$$\theta_{Y|X} = f(\theta_{Y|Z,X}, \theta_Z) \tag{15}$$

A similar relationship holds for $p_{Y|Z}$.

3.3 An overview of the different effects

Taking stock of Sections 3.1 and 3.2, we have distinguished two types of paths: a direct path, *e.g.* $Z \to Y$, and an indirect path, *e.g.* $Z \to X \to Y$. We also have distinguished several types of effects, each one being characterized by a specific conditional distribution, sometimes representing a structural sub-mechanism and sometimes not. In the saturated case, see Figure 1, one has a direct effect of a parent characterized by the final sub-mechanism and conditional distribution $p_{Y|X,Z}$, a total effect of a non-parent ancestor, *e.g.* characterized by conditional distribution $p_{Y|Z}$, a *prima facie* direct effect of a parent, characterized *e.g.* by conditional distribution $p_{Y|Z}$ or $p_{Y|X}$ and an indirect effect defined as the difference between a total and a direct effect of a given parent. Note that

in this particular saturated case, $p_{Y|Z}$ characterized both a total effect and a *prima facie* direct effect of Z on Y. The selection of the variables to be controlled for depends upon the complete specification of the recursive decomposition and upon the type of effect to be considered. In particular, one should control for the possible impact of confounding and interaction.

3.3.1 A general rule for direct effects

As a general rule, if a variable has a direct effect on an outcome, one should control for, or condition on, the other variables having a direct effect on this outcome. Only these other variables should be controlled for. Basing causality analysis on the functioning of sub-mechanisms, and in particular on the parents of an outcome, we have shown that these control variables are not restricted to confounders only but also encompass the other parents leading to variations in the distribution of the outcome variable and possibly being in interaction with the causal variable, at variance from Pearl (2000).

3.3.2 Measuring direct effects

Taking the case of Figure 1 as an example, the final sub-mechanism of interest is represented by the conditional distribution $p_{Y|X,Z}$. The *direct* effect of a variation of a causal 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. 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 assessing the effect of the variation of X from the value x to the value $x+\delta$, may be made by considering a set 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 of X on the conditional expectation may be evaluated by differentiating the conditional expectation $\frac{\partial \mathbb{E}[Y|X=x,Z=z]}{\partial x}$. As an example, suppose:

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

where α_3 measures an interaction of effects. Then

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

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.* :

$$\mathbb{E}\left[Y \mid X = x, Z = z\right] = f_1(x) + f_2(z)$$

then $\frac{\partial \mathbb{E}\left[Y|X=x,Z=z\right]}{\partial x} = f_1'(x)$ does not depend on z and is a function of x rather than a constant parameter. As already mentionned, the 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. The direct effect of X in Figures 3 and 5 is analyzed exactly in the same way whereas the case of Figure 4 is different because there Z is not a confounder anymore and is not a parent of Y.

4 Controlling in more complex models

Up to now, we have analyzed the situation of three variables, an exceedingly simple case. Once we face a more realistic situation, the issue of recursive decomposition usually becomes much more complex. As an example, Bouckaert and Mouchart (2001) analyze the impact of a drug on a disease by means of data from a clinical trial. All variables are binary. The first one codes the intake, or not, of an active product. Two variables code the presence, or not, of two different observable effects (say, a primary and a secondary effect). For each observable effect, two latent variables code the presence, or not, of the pharmacological action of the product and of the presence, or not, of a residual action that could be labelled placebo or nocebo effect. This situation gives rise to 7 binary variables, three of which are observable and four are latent. Thus a saturated model would involve $2^7 - 1 = 127$ variation-free parameters. With three observable variables, only $2^3 - 1 = 7$ parameters would have to be estimated. Imposing more structure on the saturated model is clearly crucial for obtaining an operational version of the initial model. Based on properties of boolean algebra (with binary variables) and on contextual knowledge, in this case medical knowledge of the global recursively decomposed mechanism, the authors end up with a model with 7 identified parameters that are contextually interpretable because they characterize medically plausible sub-mechanisms. These features of the modelling strategy made possible to search, in the medical literature, for prior information on the parameters used as a first stage in a Bayesian inference, as detailed in Bouckaert and Mouchart (2000).

Using the DAG terminology, the issue of control is raised here in the context of defining and measuring the effect of an "ancestor" variable (*i.e.* ancestor-parent in the case of a direct effect or ancestor-non parent in the case of a total effect), considered as a causing variable, on an outcome

variable when the global mechanism is complex, *i.e.* involving more than 3 variables. From a structural modelling point of view, and its accompanying recursive decomposition, this issue has two facets. On the one hand, 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 hand, 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 issue of control is 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. The following sections consider direct and total effects in models with more than 3 variables and propose a second general rule for determining control variables, in the case of evaluating total effects.

4.1 Direct effects in a 4-variable case

4.1.1 A saturated model

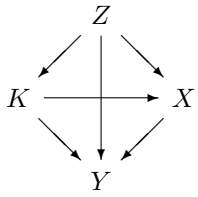

Figure 6: *A four-variable saturated case*

Figure 6 presents a 4-variable saturated model where Y is an outcome of the *direct* effect of the three other variables. The corresponding recursive decomposition can be written as:

$$p_{Z,K,X,Y} = p_Z \; p_{K|Z} \; p_{X|K,Z} \; p_{Y|K,X,Z} \tag{16}$$

Following the general rule given in Section 3.3.1, the *direct* effect of each of these three variables on Y requires controlling for the other two variables. Similarly, the *direct* effect of Z, or of K, on X, is captured by the conditional distribution $p_{X|K,Z}$ and requires controlling for the other variable, respectively K and Z.

4.1.2 An unsaturated model

Let us now consider Figure 7, a simplification of Figure 6 obtained by the following independence conditions:

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

Figure 7 corresponds to the following recursive decomposition:

$$p_{Z,K,X,Y} = p_Z \; p_{K|Z} \; p_{X|Z} \; p_{Y|K,X} \tag{17}$$

Here the sub-mechanism of interest is represented by $p_{Y|K,X}$; the direct effect of X (or of K) on Y

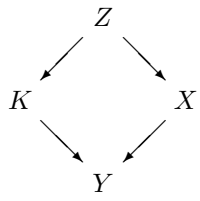

Figure 7: *A four-variable unsaturated case*

may be analyzed by measuring the effect of X (or of K) on Y for different values of K (or of X) which has to be controlled for, by application of the general rule for direct effects.

Contrary to Figure 6, Z is not a parent anymore in $p_{Y|K,X}$ but is an ancestor non-parent of Y. The total effect of Z on Y, where a variation of Z will modify the conditional distributions of the parents (X and K) of Y, may be evaluated through $p_{Y|Z}$ where:

$$p_{Y|Z} = \int\int p_{Y|K,X} \; p_{K|Z} \; p_{X|Z} \; dX \; dK \tag{18}$$

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 non-parent effect.

4.2 A 5-variable case with collider

Consider Figure 8, discussed *e.g.* in Pearl (2012). This figure corresponds to the following recursive decomposition:

$$p_{V,W,Z,X,Y} = p_V \; p_W \; p_{Z|V,W} \; p_{X|W,Z} \; p_{Y|V,Z,X} \tag{19}$$

under the following independence conditions:

$$V \perp\!\!\!\perp W \qquad X \perp\!\!\!\perp V \mid W, Z \qquad Y \perp\!\!\!\perp W \mid V, Z, X \tag{20}$$

Figure 8 may be viewed as an extension of Figure 1 obtained by adding the variables V and W. Notice, moreover, that, in Figure 1, controlling for Z is sufficient for evaluating the direct effect of X on Y whereas in Figure 8, it is not sufficient: V should also be controlled for, as an application of the general rule for selecting the control variables in the case of a direct effect. Evaluating the direct effect of each of the three parents of Y requires to control both of the other two parents.

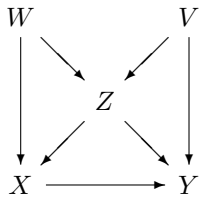

Figure 8: *A five-variate case*

Controlling only for Z may create a spurious association between V and W (indeed $V \perp\!\!\!\perp W$ does not imply $V \perp\!\!\!\perp W \mid Z$). The reader may like to compare our approach, based on the parents of an outcome, with that of Greenland, Pearl and Robins (1999) that also provides an analysis of Figure 8 on the basis of Pearl's *back-door* criterion. The latter approach leads to controlling for either Z and V *or* Z and W in order to avoid the confounding effects of Z (common cause of X and Y) and of the association between W and V induced by the control for Z. Our structural modeling approach controls for Z *and* V because they are parents of Y and takes accordingly care both of the confounding effects and of possible interaction effects between the causes of Y. Note that controlling for Z and W would take care of the confounding effects but not of the interaction effect between the two parents V and Z, at odds with the present approach.

A variable such as Z, being an outcome of two parents V and W, is called a *collider* in the graph literature. An interesting example, related to birth defects, can be found in Hernàn *et al.* (2002). To illustrate the collider issue, suppose W and V represent respectively exposure to asbestos (W) and smoking (V), Z being lung cancer, each variable being binary-coded. Assume that exposure to asbestos and smoking are independent. Then controlling for Z means examining the conditional distribution $p_{W,V|Z}$ for different values of Z, *e.g.* respectively z_1 (having lung cancer) and z_2 (not having lung cancer). It is known from the medical literature that, among those individuals having lung cancer (z_1), there is a much higher proportion of individuals having been both exposed to asbestos and smoking than among those not having lung cancer (z_2). The variables W and V are therefore associated conditionally on each value of Z. Notice that, from Figure 8, Z is an outcome variable of W and V: the distributions conditional on Z do not represent sub-mechanisms and are therefore not structural, though $p_{Z|W,V}$ is.

4.3 Controlling for latent confounders

Latent or unobserved confounders can sometimes be controlled for by proxy variables or by instrumental ones. Judea Pearl (2000) has devised a so-called *Front-Door Criterion* that can be applied for possibly controlling a latent confounder. Consider the DAG of Figure 9. To borrow Pearl's example, X could be smoking, Z tar deposit in the lungs, Y lung cancer, and U (unobserved) genotype. In

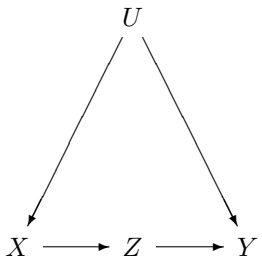

Figure 9: *Controlling for latent confounders*

this DAG, U is a latent confounder of the $X \to Z \to Y$ path, Z being an intervening variable, or mediator, between X and Y. The parents of Y are Z and U. The DAG corresponds to the joint distribution

$$p_{X,Y,Z,U} = p_U \; p_{X|U} \; p_{Z|X} \; p_{Y|Z,U} \tag{21}$$

under the independence conditions $Z \perp\!\!\!\perp U \mid X$ and $Y \perp\!\!\!\perp X \mid Z, U$. For measuring the (indirect) causal effect of X on Y, Pearl proposes a two-step procedure in order to control for the latent variable U (parentheses are ours). First, one computes the (direct) causal effect of X on Z, which is not confounded. Secondly, one computes the (direct) effect of Z on Y, which is however confounded by U as the latter is a common cause of Z (via X) and Y. But U can be controlled for by conditioning on X, which is on the path $U \to X \to Z$ and therefore "blocks" the path from U to Z (*i.e.* $Z \perp\!\!\!\perp U \mid X$). These two (direct) causal effects, of X on Z and of Z on Y, can be combined in order to yield the (indirect) causal effect of X on Y controlling for U. This procedure can be applied if the set of variables Z satisfies Pearl's *front-door criterion*: Z interrupts all directed paths from X to Y, there is no back-door path from X to Z, and all back-door paths from Z to Y are blocked by X. In Pearl's terminology, a *back-door path* is a path between two causally ordered variables that includes an arrow pointing to the first variable, such as the path $X \leftarrow U \to Y$ from X to Y in Figure 9 (see also *e.g.* Morgan and Winship, 2007).

Actually, the non-observability of U raises two difficulties. Firstly, besides Z, U is another parent of Y. Secondly, U is also a confounder, being a common cause of Z (via X) and Y. If the confounding effect of U is indeed controlled for by Pearl's procedure, it should be noticed however that variations in Y are due both to variations in Z and to variations in U. In order to obtain the direct effect of a variation in Z on Y, net of the influence of U, one would have to cancel the effect of a variation of U on Y, *i.e.* also control for U as implied by the general rule for direct effects. And this is not possible, as U is latent. More importantly, and for the same reason, possible moderator effects or interactions between the causal effects of X (via Z) on Y and of U on Y cannot be examined. Pearl's approach does not therefore completely get rid of the extraneous impact of U on the $X \to Z \to Y$ relationship, even when the front-door criterion is satisfied. In such a case, there would be no complete solution

unless the structural model be enlarged by adding new variables, such as *e.g.* parents of U that might possibly be used as proxies.

4.4 Total effects and control variable selection

Our general rule for determining the variables to control for, given in section 3.3, referred to the case of the *direct* effect of a cause X on an outcome of interest Y. We also propose a simple rule when determining the *total* effect of X on Y, composed of direct and indirect effects or more generally of multiple directed, or causal, paths from cause to outcome. The DAG represented in figure 10 is also

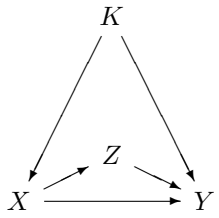

Figure 10: *Direct effects and control variables*

discussed by Pearl (2000, pp. 151-152) on the basis of his back-door criterion. In this DAG, the total effect of X on Y is transmitted by the direct path $X \to Y$ and by the indirect path $X \to Z \to Y$. As X is both a parent and a non-parent ancestor of Y, the distribution $p_{Y|X}$ characterizes a *prima facie* total effect of X on Y. However, this effect of X on Y is influenced by the confounder K, a common cause of both X and Y. There could also be an interaction (not considered by Pearl) between the effect of X on Y and that of K on Y. For these reasons, one should control for (or condition on) K, *i.e.* examine the total effect of X on Y for fixed values of K. Notice that K is not on a directed (causal) path from X to Y, though K is on a "back-door" path $X \leftarrow K \to Y$ between the two. As one sees from Figure 10, one should control for parent K of Y but not for parent Z of Y, as Z is on a directed path, in this case an indirect path, from the cause X to the outcome Y.

Consider now the more complex DAG of Figure 11 and suppose that one is interested in evaluating the total effect of X on Y through the mediators Z and K (the four variables are highlighted in bold). Here X is a non-parent ancestor of Y; the distribution $p_{Y|X}$ characterizes once again, a *prima facie* total effect of X on Y.

From the graph, one sees that the direct path $X \to Z$ is confounded by the variable L and the direct path $Z \to Y$ by the variable W. Moreover, the direct effect of M on K may be in interaction with the effect of X on K. One should therefore control for the three variables L, W, and M - parents respectively of the outcomes Z, Y, and K - by applying the general rule for determining control variables in the case of direct effects. This successive application of the rule for direct effects leads to a general rule for determining control variables when measuring the *total effect* of a variable X on

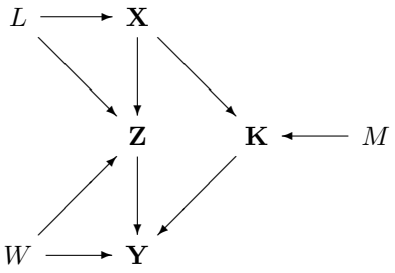

Figure 11: *Total effect and control variables*

a variable Y: *One should control for all the parents of the variables on the paths from X (excluded) to Y (included), excepting these variables on the paths themselves.* For example, one should control for the parents of Y excluding Z and K, which are on the indirect paths from X to Y, *i.e.* for variable W which is not on one of these two indirect paths. A similar reasoning applies when considering Z and K: one should respectively control for L and M, but not for X.

This suggests that the total effect of X on Y may be evaluated by comparing the conditional distributions $p_{Y|X,L=l_i,W=w_j,M=m_k}$ for a set of values (l_i, w_j, m_k), where $p_{Y|X,L,W,M}$ may be obtained as follows:

$$p_{Y|X,L,W,M} = \int \int p_{Z|X,L,W}\ p_{K|X,M}\ p_{Y|W,Z,K}\ dZ\ dK. \tag{22}$$

This is considerably more complex than the evaluation of the total effect of X on Y based on $p_{Y|X}$ that does not take into account the interaction and confounding effects of L, W and M.

5 Discussion and Conclusions

The problem of control arises when defining and measuring the effect of a cause on an outcome in a complex system. It requires a distinction between total, direct and indirect effects, either in a *prima facie* form or in a suitably controlled form. The issue of control is approached, in this paper, in a specific framework of structural modelling. More explicitly, a structural model is taken here as a representation of the underlying structure of a data generating process for a well-defined population of reference. This representation is based on background knowledge, in particular in the form of some reasonably-accepted theory and well-founded observations, *and* checked for stability, or invariance, relative to a class of interventions or of changes of environment.

In this framework, a complex, and therefore multivariate, global mechanism is recursively decomposed into an ordered sequence of sub-mechanisms represented by conditional distributions. A structurally valid recursive decomposition allows the modeller to interpret the conditioning variables as causes of the outcome. In this approach, the effect of a variation of a cause is to bring about 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 variables to be controlled for.

For an outcome of 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 from an ancestor variable to the outcome of interest.

Focusing specifically on the *outcome variables*, two general rules are proposed in this paper for deciding what variables to control for. Concerning firstly the *direct effect* of a parent X on an outcome of interest Y, control variables are those other parents of Y that have a direct effect on the outcome. Secondly, concerning the *total effect* (*e.g.* in the case of multiple causal paths) of an ancestor variable X on an outcome of interest Y, control variables are those parents of the variables on the paths from X (excluded) to Y (included), excepting these variables on the paths themselves.

The set of variables to be controlled for is thus larger than that of confounders determined by Pearl's *back-door criterion*. These rules take however exception with the statement, recalled in the Introduction, that one should control for all variables possibly having an effect on the outcome of interest, as the present approach restricts the set of variables to be controlled for to a subset of the ancestors of the outcome.

Controlling thus means examining the behaviour of conditional distributions that represent sub-mechanisms under different values of the control variable(s). These different values may however result from two different procedures. A first procedure conditions on different values of the control variable and may accordingly be called "controlling by conditioning". This operation is made on a given (structural) model and does not affect the structure of the model as it is independent of its empirical basis. This approach, based on standard rules of probability theory, is often used with observational data but could also be used with experimental data.

Another procedure intervenes in the global mechanism by fixing the values of the variable to be controlled for. It can be called "controlling by intervention". For some authors, such as D.B. Rubin and P.W. Holland (Holland, 1986), there is "no causation without manipulation", *i.e.* intervention. It should be noted that controlling by intervention implies a modification of the DGP: the variable to be controlled for is not generated anymore by the sub-mechanism identified in the recursive decomposition, but by the intervention itself. As pointed out by Lucas (1976), and often overlooked in the literature on causality, such a 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. This is not the case in the "controlling by conditioning" approach, where no changes are brought to the DGP.

An intervention that alters the DGP indicates that the mechanism is not modular, *i.e.* one cannot intervene on one part of the mechanism without altering other parts. For some, this would indicate that the mechanism is not causal. For these authors, modularity is considered a conceptual requirement for causal relations. Therefore, an important question for causal inference is *whether*,

and to what extent, a mechanism is modular, which we can only find out by looking at empirical data; for a discussion, see Illari and Russo (2014, ch.7 and 10).

Various important caveats have however to be pointed out. Some DGPs are possibly not recursive. For example, Newton's law of gravitation can be viewed as a reciprocal action among bodies, without a specific ordering that would make one body cause the attraction of the other one. Other DGPs could possibly be recursive but background information is insufficient for building the structural model and the corresponding DAG. For example, several explanations of the fertility transition[2] have been offered in the literature but there is no consensus among demographers on the correct mechanism. In this situation, checking whether one hypothesis leads to more stable mechanisms than another hypothesis may be particularly relevant but does not always provide a conclusive answer to this challenge. In other cases, a structural model can be proposed but the data are unavailable for some of the variables in the model and these remain latent. This is also the case with time-dependent feedback models, which are recursive but where data are lacking on the accurate timing of causes and outcomes.

Sub-mechanisms have to be clearly spelled out and justified, and the role-function of each variable in the sub-mechanism must be given. If there are several paths from cause to outcome, *e.g.* various risk behaviors for cardiovascular diseases, one should state whether exposed individuals can follow several paths together or whether these paths are mutually exclusive. Structural models, and the causal inferences derived from them, refer to a population and not to an individual. At the individual level, one cannot experience at the same time both the putative cause and its counterfactual, *e.g.* taking aspirin to recover from a cold and not taking aspirin. At the population level this is however not true: for colds, some individuals take aspirin and others do not, and an important issue here is to take into account the fact that the individuals in the two groups possibly differ on other factors too, populations being heterogeneous. Thus, in this paper the focus has been on conditional probability distributions instead of solely on their characteristics, such as the conditional expectation (*e.g.* the expected number of children per woman).

To conclude, there are no conditional independencies that *alone* tell us whether a variable is a control variable that has to be included in the model. Such decisions are taken, to the best of our knowledge, on the basis of background information, of preliminary analyses of data and of the structural model that is accordingly proposed. "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 and methods. In particular, the progress of knowledge can lead to uncovering previously unrecognised control variables and unrecognised mechanisms.

Acknowledgements. The authors thank Thomas Baudin, Malgorzata Mikucka and Joniada Milla,

[2]See for instance a special issue of *Population and Development Review* (2015) on the fertility transition.

for valuable comments and suggestions on an earlier version of this paper.

Appendix : DAG and Recursive decomposition

The graph illustrated by Figure 11 illustrates a case of non-uniqueness of a recursive decomposition corresponding to a given DAG. Indeed, this DAG may correspond either to the recursive decomposition:

$$p_L\, p_W\, p_M\, p_{X|L}\, p_{Z|X,L,W}\, p_{K|X,M}\, p_{Y|W,Z,K} \tag{23}$$

or to:

$$p_L\, p_W\, p_M\, p_{X|L}\, p_{K|X,M}\, p_{Z|X,L,W}\, p_{Y|W,Z,K} \tag{24}$$

In both cases, these recursive decompositions incorporate the following independence hypotheses:

$$L \perp\!\!\!\perp W \perp\!\!\!\perp M \qquad X \perp\!\!\!\perp W, M \mid L \qquad Y \perp\!\!\!\perp L, X, M \mid W, Z, K \tag{25}$$

But the first one also incorporates:

$$K \perp\!\!\!\perp Z, L, W \mid X, M \qquad Z \perp\!\!\!\perp M \mid X, L, W \tag{26}$$

whereas the second one assumes

$$K \perp\!\!\!\perp L, W \mid X, M \qquad Z \perp\!\!\!\perp K, M \mid X, L, W \tag{27}$$

Equations 23 and 24 represent a same multivariate distribution: the two recursive decompositions are therefore observationally equivalent. But the hypotheses 26 and 27 bear on different conditional distributions and thus on different sub-mechanisms. Thus, for instance, hypothesis 26 assumes that the conditional distribution $p_{K|Z,L,W,X,M}$ does not depend on Z, L, W whereas hypothesis 27 assumes that the conditional distribution $p_{K|L,W,X,M}$ does not depend on L, W. This example shows that background knowledge may have a crucial role in deciding which conditional distribution is actually structural and which independence conditions should be tested. The model-builder should first order all the variables, as a basis for a recursive decomposition. Next, for each factor, the conditional independence conditions should be elucidated and finally the DAG proposed with the arrows corresponding only to the cause-effect relations.

Figure 11 and decompositions (23) and (24) furthermore raise questions about the ordering of the variables once the independence conditions have been incorporated. Indeed, Figure 11 leads to the following hierarchy of the variables:

$$\{L, W, M\} \prec \{X\} \prec \{K, Z\} \prec \{Y\} \tag{28}$$

with two equivalence classes: $\{L, W, M\}$ and $\{K, Z\}$, where $\{\alpha\} \prec \{\beta\}$ should be read " $\{\alpha\}$ precedes $\{\beta\}$". To the first equivalence class correspond $3! = 6$ recursive decompositions corresponding

to different orders inside this equivalence class and therefore to different possible sub-mechanisms. Although these sub-mechanisms may have different plausibilities as viewed from background knowledge, these 6 representations correspond to a same DGP. A similar remark holds for the second equivalence class, with the additional remark that the 2 orderings require the same variables to be controlled for.

References

Atinc G., Simmering M.J., Kroll M.J. (2012), Control variable use and reporting in macro and micro management research, *Organizational Research Methods*, **15**(1), 57-74.

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

Baron, R. M. and Kenny, D. A. (1986), The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations, *Journal of Personality and Social Psychology*, **51**, 1173-1182.

Becker T.E. (2005), Potential problems in the statistical control of variables in organizational research: a qualitative analysis with recommendations, *Organizational Research Methods*, **8**(3), 274-289.

Blossfeld, H.-P. (2009). Causation as a generative process, Chapter 5 in: Engelhardt, H., Kohler, H.-P. and Fürnkranz-Prskawetz, A. (eds.), *Causal analysis in population studies*, Springer, 83-109.

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

Bouckaert, A. and Mouchart, M. (2000), SORE Modelling for Clinical Trials: A Bayesian Perspective, *Revista de la Sociedad Chilena de Estadística (SOCHE)*, **16-17**, 7-28.

Bouckaert, A. and Mouchart, M. (2001), Sure Outcomes of Random Events: a model for clinical trials, *Statistics in Medicine*, **20**, 521-543.

Cox, D.R. (1992). Causality: some statistical aspects, *Journal of the Royal Statistical Society*, Series A, 155(2), 291-301.

Crossman, A. (2014). Control variable, at: `http://sociology.about.com`, accessed 08/26/2014.

Fagiolo G., Moneta A. and Windrum P. (2007), A Critical Guide to Empirical Validation of Agent-Based Models in Economics: Methodologies, Procedures, and Open Problems, *Computational Economics*, **30**, 195-226.

Fisher R. A. (1935, 1st edition), *The design of experiments*, Edinburgh: Oliver & Boyd.

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

Glossary of Statistical Terms (2014), Control, at: `http://www.stat.berkeley.edu`, accessed 08/26/ 2014.

Greenland S., Pearl J., and Robins J.M. (1999), Causal diagrams for epidemiological research, *Epidemiology*, **10**(1), 37-48.

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.

Holland P.W. (1986), Statistics and Causal Inference, *Journal of the American Statistical Association*, **81**(396), 945-960.

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

Illari, P. and Russo, F. (2014) *Causality: Philosophical Theory Meets Scientific Practice*, Oxford University Press, Oxford.

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.

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 Foundation Monograph 10, John Wiley, New-York.

Lucas, R. (1976), Econometric Policy Evaluation: A Critique, in Bruner K. and Metzler, A. *The Phillips Curve and Labour Markets*, Carnegie-Rochester Conference Series on Public Policy, **1**, New York: American Elsevier, 19-46.

Mill, J. S. (1843), *A System of Logic, Ratiocinative and Inductive: Being a Connected View of the Principles of Evidence and the Methods of Scientific Investigation*, (Seventh (1868) edn), Longmans, Green, Reader, and Dyer, London.

Morgan S.L. and Winship C. (2007), *Counterfactuals and Causal Inference*, Cambridge University Press, New York.

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

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

Mouchart M., Russo, F. and Wunsch, G. (2011), Structural modelling, Exogeneity and Causality, Chapter 4 in: Engelhardt, H., Kohler, H.-P. and Fürnkranz-Prskawetz, A. (eds.), *Causal analysis in population studies*, Springer, 59-82.

Pearl J. (1988), *Probabilistic reasoning in intelligent systems: networks of plausible inference*, Morgan Kaufmann, San Mateo.

Pearl J. (2000), *Causality. Models, Reasoning, and Inference*, Cambridge University Press, Cambridge, revised and enlarged in 2009.

Pearl J. (2012), The Mediation Formula: A guide to the assessment of causal pathways in nonlinear models, in C. Berzuini, A. P. Dawid and L. Bernardinelli (eds.) *Causality: Statistical Perspectives and Applications*, Wiley, Chichester, 151-179.

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.

Russo, F.(2014), What invariance is and how to test for it, *International Studies in Philosophy of Science*, **28**(2), 157-183, DOI: 10.1080/02698595.2014.932528

Schjoedt L. and Bird B. (2014), Control variables: use, misuse and recommended use, in Carsrud A. and Brännback M. (eds.) *Handbook of Research Methods and Applications in Entrepreneurship and Small Business*, Edward Elgar, Cheltenham, 136-155.

Spector P. and Brannick M.K. (2011), Methodological urban legends: the misuse of statistical control variables, *Organizational Research Methods*, **14**(2), 287-305.

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. (2014), Functions and mechanisms in structural-modelling explanations, *Journal for General Philosophy of Science*, **45**, 187-208.