Causality in the Social Sciences: a structural modelling framework

Russo, F.; Wunsch, G.; Mouchart, M.

DOI
10.1007/s11135-019-00872-y

Publication date
2019

Document Version
Final published version

Published in
Quality and Quantity

Quality & Quantity (2019) 53:2575–2588
https://doi.org/10.1007/s11135-019-00872-y

Causality in the Social Sciences: a structural modelling framework

Federica Russo[1] · **Guillaume Wunsch**[2] · **Michel Mouchart**[3]

Published online: 17 April 2019

Abstract
There is no unified theory of causality in the sciences and in philosophy. In this paper, we focus on a particular framework, called structural causal modelling (SCM), as *one* possible perspective in quantitative social science research. We explain how this methodology provides a fruitful basis for causal analysis in social research, for hypothesising, modelling, and testing explanatory mechanisms. This framework is not based on a system of equations, but on an analysis of multivariate distributions. In particular, the modelling stage is essentially distribution-free. Adopting an SCM approach means endorsing a particular view on modelling in general (the hypothetico-deductive methodology), and a specific stance on exogeneity (namely as a condition of separability of inference), on the one hand, and in interpreting marginal–conditional decompositions (namely as mechanisms), on the other hand.

Keywords Structural causal modelling · Recursive decomposition · Mechanisms · Causality - Causal modelling

1 Introduction

There is no unified theory of causality in the sciences and in philosophy. Many accounts of causality have been developed and no consensus has been reached so far. In their monograph on the philosophy of causality, Illari and Russo (2014) developed a pluralistic approach, according to which causal theory resembles a *mosaic*, in which different tiles—appropriately placed—let an image emerge. Relevant tiles for a causal theory correspond to different philosophical questions or scientific problems one addresses, and the accounts thus far developed—for instance in terms of mechanisms, processes, or counterfactuals—help shed light on the different approaches to causal analysis adopted in the sciences. The

✉ Michel Mouchart
michel.mouchart@uclouvain.be

[1] Department of Philosophy, University of Amsterdam, Amsterdam, The Netherlands

[2] Center for Demographic Research, University of Louvain (UCLouvain), Louvain-la-Neuve, Belgium

[3] Institute of Statistics, Biostatistics and Actuarial Sciences (ISBA), University of Louvain (UCLouvain), Louvain-la-Neuve, Belgium

social sciences are no exception in this respect, and a suitable causal mosaic is needed in this area too. The suggestion has been promptly taken up by e.g. Johnson et al. (2017), who explain how, in the specific area of mixed methods, the causal mosaic approach well accounts for the different methods, concepts, and assumptions that are at stake.

In this paper, we focus on a particular framework, called structural causal modelling (SCM), as *one* possible perspective in quantitative social science research and explain how this methodology provides a sound basis for causal analysis in social research. This is indeed required if one wishes to go beyond description in quantitative social science, with the purpose of explaining correlations among variables or of developing relevant policies in the social realm. Our modelling approach is thus part of the larger mosaic of causal methods available in social research. In a mosaic framework, different conceptual and methodological aspects are at stake. This is an important form of pluralism that contributes to turning a statistical model into a causal model.

"Structural modelling" is a label used by a number of methodologists, practicing scientists, and philosophers of science in the contexts of quantitative social science, economics and econometrics, and in computer science. The objective of this paper is to present an overview of the structural causal modelling framework developed by the present authors in a series of articles spanning the last decade or so. In previous contributions we analysed specific aspects of our approach to structural modelling, and here we wish to discuss the origin of the approach and its main characteristics. We aim to offer practicing scientists an operational tool when choosing and using structural modelling.

We first introduce our approach with an example drawn from the field of demography. Consider a study on the recourse to contraception in urban Africa, the example being taken from Gourbin et al. (2017). The countries are characterized by different levels of contraceptive prevalence, but also by the different effectiveness of the methods used. Several questions may be raised, the following two amongst others. Firstly, what is the hierarchical ordering of causal relationships among the individual factors involved in the use of contraception in the urban populations considered? Secondly, as education is a major factor of fertility transition, are two main indirect pathways that have been proposed in the literature (a union-reproductive path and a socio-cultural one), leading from woman's education to contraceptive use, confirmed by the data?

Most analyses of contraceptive use have had resort to statistical methods that do not take into account a possible causal ordering among the variables, implicitly assuming that all the putative determinants just have a direct effect on the dependent variable. However, the impact of these various factors on the use of contraception can be direct or indirect, meaning in the latter case that the effect of some putative causes can be mediated by one or more intermediate factors. To answer the questions raised above, one needs an approach that allows researchers to propose an *explanatory mechanism* for the outcome of interest, composed of various sub-mechanisms, and subsuming in particular the distinction between mediators, moderators, and confounding variables.

The purpose of this paper is to explain how the SCM framework provides the tools to hypothesise, model, and test explanatory mechanisms. Our approach proves particularly useful in social science contexts, since it is well suited to observational data; in fact, in social science experiments are notoriously difficult to carry out for ethical or practical reasons, and our approach allows social scientists to go beyond mere description and to propose a causal explanation even in absence of experiments and interventions.

The paper is structured as follows. In Sect. 2 we present the origins of the structural approach, thus showing that the SCM framework is rooted in well-established traditions in social science methodology. Yet, we positively contribute to improving on this solid basis

by developing specific methodological and epistemological aspects, further explained in Sect. 3. The latter provides an overview of SCM, starting with its methodological foundations, especially the hypothetico-deductive methodology and the characteristics that a statistical model ought to include. We give particular emphasis to the marginal–conditional decomposition because, when given a suitable interpretation as a mechanism, it allows us to explain a phenomenon of interest. We next turn to important epistemological aspects of structural modelling, notably the role and use of background knowledge, and the important role of the notion of variation in the modelling process. The section further discusses how structural causal models aim to build generic knowledge, in order to explain cause-effect relations and to inform policy. We compare our approach with the widespread methodology of potential outcome models. Section 4 further discusses other aspects related to structural modelling, notably the role of agents, of data and of time. In the concluding section we offer a general reflection on the aims of structural causal models.

2 A dual origin

The structural approach to causation, as summarized in Sect. 3, 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 models derived from economic theory. The recursivity of a model was shown to facilitate statistical inference. Wold (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, at a given stage of decision. In this approach, the ordering of the endogenous variables is given a priori by economic theory. Simultaneity of equations may moreover be viewed as an effect of incomplete observability, in particular of time aggregation. More 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 et al. 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— as proposed by Blalock, Duncan, and Boudon among others. The purpose here is to express a correlation between putative causes and effects through a recursive decomposition of the joint distribution of the variables. These authors represented the relations in branching sequential order by an arrow diagram, a graphic representation taken up later and expanded by Judea Pearl especially, with his directed acyclic graphs (DAG) approach to causality (Pearl 2000).

The causal modelling framework presented in Sect. 3 is however substantially different from the approach developed in the literatures on simultaneous equation models in econometrics and on structural equation models in the social sciences. In particular, our framework is not based on a system of equations, but on an analysis of multivariate distributions. Moreover, the modelling stage is essentially distribution-free. The distributional hypotheses are accordingly introduced only at the final stage of inference. These ideas will be thoroughly explained later in Sects. 3 and 4. More substantially, the usual structural approaches endeavor at *representing* causal knowledge using systems of equations, while the structural perspective presented here aims at *constructing* causal knowledge by hypothesizing,

modelling, and testing explanatory mechanisms for the phenomenon of interest. Thus the emphasis is placed on mechanisms and sub-mechanisms, on the role-function of variables in these sub-mechanisms, on background knowledge, and on structural stability. Section 3 provides an overview of the approach and Sect. 4 develops further the major characteristics of the present model, compared to those in economics or in social science.

3 A structural causal modelling framework

This section provides an overview of the SCM framework. We explain the methodological and epistemological foundations in Sects. 3.1 and 3.2 respectively. Sections 3.3 and 3.4 discuss how structural causal models aim to build generic knowledge about a phenomenon of interest for explanatory purposes and for informing policy making.

3.1 Methodology of structural causal modelling

The structural causal modelling approach may be viewed as a chapter in the domain of statistical modelling, where a statistical model is considered as a set of "reasonable" hypotheses concerning the *data generating process* (DGP) represented as a probability distribution. A stochastic representation of the DGP is used to explain a phenomenon of interest and such explanation involves two elements: (1) a stochastic element that embodies what is not explained in the working of the DGP (i.e. of the mechanism) and (2) a non-stochastic element, the characteristics- or parameters- of the probability distribution, that provides the nature of what is explained by the statistical model. Said differently, the statistical model provides only a partial explanation of the mechanism of the DGP.

Compared to structural models in econometrics or in social sciences, the present structural framework takes distance from the latter in several aspects that we outline below. To begin with, our structural approach is based on a *hypothetico-deductive* (*H–D*) *methodology*. This means that a hypothesis is first formulated, a model developed and tested, and the results interpreted in order to confirm or disconfirm the initial hypothesis. H–D methodologies are widely used in science and are often associated with the falsificationist view of Karl Popper. However, in philosophy of science, hypothetico-deductivism has been developed much beyond the original Popperian approach detailing, among others, the role of background knowledge at the hypothesis formulation stage or the fact that we learn also from disconfirmed hypotheses—so models can be iteratively improved on, and we don't start each time from scratch. We now discuss other important methodological features.

Causal and structural. Focused on causal analysis, this approach depends upon reliable background information and evidence for assessing causes of outcomes and evaluating effects of causes, and more generally on the structure of relations between causes and outcomes.

Recursive decomposition and DAG. 'Explaining' essentially means here representing and decomposing a complex and global mechanism in terms of a set of simpler sub-mechanisms. The explanation is based on a recursive decomposition of the joint distribution of the variables entering the statistical analysis. This recursive decomposition is equivalent to a systematic marginal–conditional decomposition according to a specific ordering of the variables. Thus the joint distribution is written as a product of conditional distributions

where the conditioning variables form an increasing sequence and where each factor of this product represents a sub-mechanism. For that reason, directed acyclic graphs (DAGs) provide a privileged tool of representation, though a DAG does not allow representing all particularities of a multivariate distribution nor of a recursive decomposition.

Exogeneity and causation. Under a suitable exogeneity condition of non-confounding, a confounder being a common factor of cause and outcome, the conditioning variables are viewed as causing variables in the sub-mechanism where they appear. This is the reason why the structural model is also called a causal model, because causation is relative to a particular model built with the purpose of eliciting causes.

Distributions rather than equations. The basic objects of analysis are sets (in product form) of distributions rather than sets of equations. There are two reasons for that. Firstly, a (structural) equation with random variables may have an ambiguous meaning because of mismanaged conditions of exogeneity. For instance, discussions about normalization of equations, in the econometric literature of simultaneous equations, have been obscured by considerations of identification. Secondly, equations are related, at best, to conditional expectations, although effects of causes may take other ways. For instance, in actuarial applications, the effect of some contracts may be more in the tails of the distributions than in the expectations.

Explanation and parametrization. Explanation is based on a recursive decomposition, the latter taken as the first stage of explanation. As mentioned before, representing a DGP by a probability distribution implies that this representation leaves unexplained some part of the DGP, namely the stochastic component of the model. Therefore, the statistical explanation concerns the characteristics, or parameters, of the probability distributions. This fact raises the issue of the specification of the parametrization. Mathematically, the parametrization of a family of distributions may be viewed as a labeling system where the value of the parameter identifies a particular distribution in the whole family. For this role, the parametrization is arbitrary. For instance, it is inconsequential, as far as labeling the normal family is concerned, whether the normal family is parametrized by means of the expectation and the variance, or the expectation and the standard deviation, or the inverse of the coefficient of variation and the inverse of the variance (as is natural from the point of view of the exponential family). Once a conditional distribution is deemed to represent a specific sub-mechanism, the role of the parametrization is also to identify the operation of the sub-mechanism (more information in Mouchart and Orsi 2016).

Stability and invariance. Considering as structural a mechanism underlying the workings of a DGP requires that the model enjoy suitable properties of stability, or invariance, under an adequate class of interventions and of modifications of the environment. Indeed, a model that would be different for, say, each observation should not be considered as structural. Said differently, the issue here is to look for a proper separation between the incidental and the structural aspects of the DGP. From a statistical point of view, this issue is also that of the population of reference. One reason for this is that no model in the social sciences can pretend to be universal in time and in space. It should be stressed that this stability, or invariance, regards both the recursive decomposition, and therefore the ordering of the variables, and the value of the parameters.

3.2 Epistemology of structural causal modelling

Modelling in science is always accompanied by specific epistemological standpoints, views, or concepts. SCM is no exception in this respect. An epistemology for the structural modelling framework ought to shed light on how we reason during the whole H–D process. We focus here on three selected aspects of this epistemology: the meaning and role of background knowledge, the goal of a structural model to explain a phenomenon, and the role of the notion of variation in establishing causal relations.

Background knowledge. Background knowledge plays a crucial role at each stage of the H–D methodology. Firstly, causal attribution is often quite a difficult issue once a system becomes complex, as is often the case in social science. Secondly, background knowledge typically involves theories concerning the domain of analysis, but also embraces a much wider scope, in particular involving previous results and preliminary analysis of data. It is on this basis that a preliminary hypothesis is formulated. Background knowledge is likewise involved in the process of developing a specific statistical model, where one makes important choices about the parametrization, testing methods, etc. Again, the results of tests are interpreted against available background knowledge. This means, effectively, that any result achieved *is* dependent on what we already know, and yet, with structural modelling we also aim to go *beyond* what we already know. This introduced a well-known tension in scientific research, namely that there is no theory-free genuine discovery (*pace* the staunchest advocates of big data practices) and that to assess the genuine novelty of a scientific result is far from being an easy task. The SCM approach does not aim to be a magic recipe for solving this evergreen conundrum but aims to help researchers being aware of their tacit or explicit use of background knowledge during the whole process.

Representing and explaining. In this approach, a model aims not only at representing the DGP in probabilistic terms but also at explaining the functioning of a complex mechanism, the phenomenon of interest. An important methodological feature—mentioned earlier in Sect. 3.1—is the recursive decomposition. This decomposition does more than just representing a phenomenon probabilistically: it aims at explaining its (probabilistic) *structure*. Thus the marginal–conditional decomposition says which variables are the causes and which are the outcomes, and for complex systems with more than three variables, it details the whole organization of more proximal and distal causes. Whether variables figure as causes or as effects depends partly on our background knowledge (see above) and partly on preliminary data analysis and tests. The explicit interpretation of independent and dependent variables as causes and outcomes is part of a broader view according to which the marginal–conditional decomposition represents a *mechanism* underlying the phenomenon of interest. In the social science contexts, these will be i.a. social, economic, psychological mechanisms.

Variation. With structural modelling we aim at explaining a phenomenon and we do so by detailing a marginal–conditional decomposition that effectively provides a specific structure in which causes and effects are organized. At the level of testing, what we are interested in is checking whether—and to what extent—variations in the putative cause(s) are accompanied with variations in the putative effect(s). The notion of variation is at the very basis of a causal epistemology for structural modelling for a number of reasons (Russo 2009). To begin with, a phenomenon becomes of interest when unexpected, exceptional,

or surprising variations are found out. For example, contraceptive levels are not the same among different groups—why is it so? Contraceptive methods differ as to their effectiveness—why is it so? These could be two questions at the basis of the example provided in the introduction. Next, causal methodologies are essentially based on the idea that we need to compare cases—we test what varies with what, in which case, etc. The importance of the notion is not just conceptual (understanding what it actually means to test for causal relations) but also methodological. In fact, if tests for causal relations essentially concern variations within and between groups or characteristics, it is of utmost importance to define the population of reference, its sub-groups, the variables to be measured and the way they are measured. It may make a difference whether we stratify the group of women in bigger or smaller age groups, as some variations may emerge while others do not. Likewise, the way a variable is measured may or may not make explicit the threshold at which a cause becomes active.

3.3 Constructing generic knowledge for explanatory purposes

The SCM framework is focused on general (or type-level) causation, as distinct from singular (or token-level) causation: it aims at deriving general causal claims for a population of reference. More specifically, structural models *construct* generic causal knowledge by modelling mechanisms (in the sense explained above). To explain the relevance and importance of this point, we contrast and compare our structural causal modelling approach with the widespread approach in terms of counterfactuals, also known as the 'potential outcome model'.

To illustrate our point, we use a stock example from the literature, namely whether aspirin[1] (A) relieves headache (H). Other more relevant examples will be drawn from the field of demography. In the counterfactual approach or potential outcomes framework, as considered by Donald Rubin, James Heckman, and many other researchers in the field of social sciences (see e.g. Morgan and Winship 2007), one compares the factual situation (e.g. taking aspirin A) to the counterfactual situation (not taking aspirin) in a given *reference population* suffering from headache H. Then one checks for those taking aspirin, the putative cause of interest, if taking aspirin makes a difference in the evolution of the headache, compared to those not taking aspirin. Using binary variables, is the probability of H dissipating (i.e. changing from $H=1$ to $H=0$), during a specified period of time, different for those taking aspirin ($A=1$) and for those not taking aspirin ($A=0$)? In other words, does the variation in the cause A bring about a variation in the outcome H? (here the importance of the notion of variation—presented above—becomes apparent). The potential outcomes framework is derived from the well-established practice of randomized trials, where the effect of a drug is obtained by comparing a treatment group taking the drug to a control group not taking the drug, individuals being assigned at random to the groups. The potential outcomes framework is however put into use in observational studies where randomized assignment has not been applied for ethical or practical reasons.

Of course, the counterfactual does not make sense at the individual level, as $A=1$ and $A=0$ cannot occur at the same time: one cannot take and not take an aspirin simultaneously. Furthermore, as pointed out in Russo et al. (2011), the counterfactual $A=0$ does not cover all the possible differences in behaviors, compared to the group taking aspirin, that

[1] Acetylsalicylic acid. The name 'Aspirin' is a Bayer brand.

could have an impact on H. One needs, in addition, to take into account these other possible differences between the individuals taking aspirin and those not taking aspirin that could have an effect on H. Specifically, one must *control* for possible *confounders*, i.e. common causes of A and H, by conditioning on all the other so-called *direct causes* of H, as proposed e.g. in Mouchart et al. (2016). If the cause of interest is *conjunctive*, one excludes however from control the other components of this conjunctive cause (Wunsch et al. 2018), as the variation of all components would be needed here for the outcome to occur.

From the above, one can add that a counterfactual approach does not make sense at the individual level, and that it requires controlling for possible confounders at the population level. When these confounding variables are taken into account, one can then compare the group where A = 1 to the group where A = 0, *i.e.* the group where there is a change in A from 0 to 1 to the group where there is no variation in A, and observe if taking aspirin makes a difference in H. If there is a difference, the variable A after proper control is deemed to be *exogenous* and can be considered having a causal impact on H (Mouchart et al. 2009). Actually, one should not rely on a sole study, but the research should be replicated to check whether the result is *invariant* and can be generalized to the reference population and period considered. For example, in a study on the impact of smoking on cardiovascular diseases (CVD) in Europe, one should test if this risk factor has the same effect in e.g. Portugal, Finland, and Germany. If it unexpectedly does not, taking the whole of Europe as the reference population would not be adequate, and further research would be needed for understanding why the effect of smoking on CVD differs among these populations; for instance, have all relevant confounders been controlled for?

But what are the *direct causes* of H other than A? The answer depends upon one's *background knowledge* concerning the other proximate variables or direct causes having a possible impact on the variation of H. The change in H could not only be due to the change in A but also to changes in these other variables X,Y, Z…, and the latter have therefore to be controlled for, i.e. conditioned on. This also takes account of possible *interaction effects* among causes (Mouchart et al. 2016). Moreover, these other variables are *context-dependent*: in another population of reference, the set of variables would possibly not be the same. For example, the direct causes of H other than A might differ between blue collar workers and white collar ones, due to different work environments and behaviors. One has thus moved from a simple case with two variables only to a more complex multivariate causal situation. More generally, a fuller model would also examine the biological mechanisms leading from A to H.

3.4 Constructing general knowledge for policy purposes

Pursuing our illustrative example, for scientific knowledge and for health policy purposes, it is useful understanding why some individuals take aspirin and others do not, when they have a headache. In other words, what are the direct causes of the direct cause A of H? And, if relevant for the purpose of the study, what are the direct causes of the causes of the cause A, etc. More generally, what are the causes of the various direct causes of H, i.e. A, X, Y, Z, …? Once again, background knowledge may suggest the answers for a given population of reference and period, if these claims are empirically verified. Graphically, the causal model thus derived, in this 'bottom-up' process, corresponds to a DAG or *directed acyclic graph* (Pearl 2000) of which the final outcome is H. Note that the same DAG can be obtained in a 'top-down' approach by *recursively decomposing* the joint distribution

of all variables considered to have a direct or indirect effect on H (Mouchart et al. 2010). This requires that the DAG be unique, i.e. that the set of cause-effect relations derived from background knowledge be unique. If this is not the case, different causal models could lead to the same final outcome H. Thus it is important to correctly specifying, to the best of our knowledge, the DAG and its underlying causal framework. Once again, the study should be replicated to see if the results are *invariant* and can be generalized for the reference population and period considered.

The present approach can deal with various types of variables, such as interval (e.g. income), ordinal (e.g. social status) or categorical (e.g. religion). What counts, among others, is the effect a variation of an exogenous cause can have on an outcome variable, for example the impact on fertility of differences in income or in social status, or of having a different religion, after due control for confounders.

Constructing the DAG and its underlying causal model requires taking into account all available evidence and not only paying lip-service to one or another established theory. In particular, one should not only consider results of quantitative research but also information provided by qualitative studies of small target samples drawn from the population of reference. SCM should however not be data-driven, in the sense that—at the conceptual level—the causal model should be in dialogue but not dependent upon the actual data to be analysed. Nevertheless, the lack of suitable data for testing the model is a problem frequently encountered in practice. In other words, the available evidence leads to a conceptual framework that cannot be completely reproduced by the operational framework expressed in terms of the data at one's disposal. For example, in the contraceptive study referred to above, though the data available did allow a relatively appropriate description of the union-reproductive path, the same could not be said for the socio-cultural path, because the variables used as indicators of gender relationships and of social and economic capital were far from allowing a good operationalization of these concepts.

The establishment of the set of direct and indirect causal paths leading to the final outcome (e.g. H) is based on the putative knowledge of the *mechanism* and *sub-mechanisms* producing the various causes and effects. The DAG is thus a representation of the data generating process (DGP) producing H. The causes are incorporated on the basis of their *function* or *role* in the mechanism (Mouchart and Russo 2011; Wunsch et al. 2014). In other words, role-functions are the reasons why some variables and relations are chosen rather than others. This is indeed what the recursive decomposition is intended to do, i.e. to model the DGP in terms of mechanisms. Social mechanisms are of course unobserved analytical constructs that propose hypothetical causal links between observed events. We can however test mechanisms via their predicted effects, i.e. by comparing implications of the mechanisms with empirical facts. Nevertheless, the progress of knowledge can eventually lead to proposing other mechanisms for the DGP: there are no absolute truths in science.

Recursive systems, such as the one presented here, require a unidirectional relationship among variables, with no simultaneous feedback. It often happens however that a complete ordering of variables cannot be accomplished, either because some causes occur jointly, as contingent conditions or as conjunctive causes acting together on the outcome, or because one does not have the information required for determining causal priority among the variables. In these cases, the recursive decomposition is not complete anymore, in the sense that some of the conditional distributions are not univariate but concern vectors—or 'blocks'—of variables. We have nevertheless shown (Wunsch et al. 2018) that in many such cases a preliminary, though often imperfect, causal analysis can be performed on the basis of a detailed study of each context, including contingent conditions, and/or of possible additional evidence.

Up to now, the structural model proposed here is static: time is not incorporated explicitly into the framework. Specifically, reverse causation and feedback effects are not explicitly taken into account. The structural modelling perspective has been applied, in particular, to the study of the direct and indirect paths leading to contraceptive use in urban Africa, the issue raised in the introduction to this paper (Gourbin et al. 2017). However, the individuals concerned were interviewed at a same point in time and not over their lifetime. Results refer therefore to inter-individual differences obtained at the time of the surveys providing the data, and not to life-course differentials as could be derived from retrospective or prospective data.

4 Agents, data, and time

In the social sciences, the basic actors or agents are individuals. Their behavior determines, among others, the levels and trends in mortality, fertility and migration. For example, family decisions on fertility are taken by interacting individuals. Institutions play a role, undoubtedly, but they are run, till the present time at least, by human beings, though this could change in the future, taking into account the huge current progress in artificial intelligence, including *deep learning*. The structural modelling perspective presented here is thus inserted in the framework of *methodological individualism*, i.e. the fact, following Max Weber, that social phenomena result from the motivations and actions of individual agents, and from their interactions. The basic statistical units of structural models are therefore individuals, the perspective being rooted at the micro level.

This does not, of course, exclude the fact that the institutional and socio-cultural contexts play an important role in the motivations, intentions and behaviors of the individuals, making the structural model context-dependent. If these macro-level factors are constant *contingent conditions*, such as the provisions of constitutional law applicable to all, they can be incorporated into the description of the reference population and not into the model itself, as proposed by Wunsch et al. (2018). In this sense, our approach comes close to what Little (2006) has called 'methodological localism', which aims to emphasise how individual agents are always socially situated and their actions occur therefore in a particular context. More generally, contingent macro conditions reflect a state of stability, or absence of change, of variables that are necessary (but not sufficient) for the working of the causal (sub)mechanism. The concept of stability is relative to a specific context or field. Within this context, a stable variable should however not be considered as a cause. For this reason, a contingent stable condition is not strictly viewed as a cause, as by definition it does not vary, though it is part of the so-called "full cause", in Mackie's (1965) terms, of the outcome.

If the macro-level factors vary, on the other hand, they should be incorporated into the structural model itself. For example, if social security benefits in the event of sickness vary according to one's professional status, say being employee or self-employed, this should be taken into account in the model. If, on the other hand, the benefits are applicable to all in the same manner, they can be left as a (invariant) characteristic of the reference population. The latter should nevertheless be thoroughly described, as a part of the explanation of the phenomenon studied. For example if one compares, for one or another reason, contraceptive behaviors in Sweden and in Chad, the socio-cultural contexts of both countries should be carefully described, as they are particularly relevant for the topic studied. In other words, the contingent conditions should be studied as much as the causes. Causal

models in the social sciences are actually valid only for specific contexts and time-periods, and the context and period should always be explicated.

As the structural model presented here aims at representing causal relations among variables, the latter should be ordered according to their causal priority, which implies *inter alia* a temporal ordering of causes and effects. For example, in the field of internal migration, if one is interested in the relation between changes in occupation and changes in residence, one should know if the change in occupation (or the decision to change) has preceded the change in residence, or vice-versa. The sources of data should therefore give the time-ordering of events, in a life-course approach. This can be accomplished, for example, by retrospective or prospective population surveys, by specific registries and other forms of permanent registration of individual events, such as a national register, etc.

Many sources of data actually refer to the same individuals. A same individual may for example be interviewed in a health survey, and later on recorded in a heart disease register, with information collected by the health insurance system and hospital statistics, and end up with his/her causes of death specified on a death certificate. If each individual receives at birth a personal identification number (PIN), data from multiple sources can be linked together. One can thus examine for an individual e.g. the move from good health to ill health, then to chronic disease, disability, and finally to death, possibly also taking into account various characteristics of the individual (such as education and employment) and their change over time. One must however consider the fact that an event is often the result of a temporally prior decision-making process, based on the preferences, values, emotions and beliefs of the agents in possible interaction with others, as in the occupation/migration example referred to above. Data on the decision-making process are unfortunately most often unavailable. Contrary to the time-ordering of events, that of the various decision-making processes is thus difficult to specify.

If individual longitudinal data are available, the causal model can also easily take into account reverse causation and feedback effects, by time-ordering the variables. For example, one may suspend contraceptive use in order to have a child, and then resume contraception afterwards in order to avoid further childbearing. An elementary DAG of the model in this case is $C_t \rightarrow F_{t+y} \rightarrow C_{t+ny}$ where C stands for contraceptive use and F for fertility, the time indicators being ordered as $t < t+y < t+ny$. The problem here is an issue of data availability and not of modelling.

The situation becomes more complex when one considers the possible nature of the variables. As pointed out in Sect. 3, SCM can deal with various types of variables, interval, ordinal, or categorical. However, a variable can be an indicator of an event, such as a change in occupation, or of a process, such as smoking behaviour over time. Similarly, the outcome can be an event, such as a change in residence following a change in occupation, or a process, such as the progressive development of coronary heart disease due to smoking. The causal model should be able to deal with both events and processes, in the latter case for example by specifying for each individual the duration of exposure to smoking. Going back to our illustrative example, taking an aspirin is an event but its impact on the headache is usually gradual and not all at once. The outcome variable should thus take duration into account, by specifying if the headache is dissipated after 5, 10, … minutes after taking the aspirin, or not at all. More generally, the time-patterns of the cause and outcome variables should be taken into account.

Unfortunately, one is often confronted with an inadequate observation of the temporal sequence of events. This is typically the case when events are not recorded, or made available, on a continuous time-basis but by discrete periods of time. Depending upon the length of the observation period, it is not always possible to see if the putative cause actually

occurred before or after the outcome, and thus to determine causal priority. As discussed in Wunsch et al. (2010), time can sometimes be embedded however in background knowledge and, in this case, the causal ordering of variables can approximately be determined. For example, from cross-sectional survey data one can study fertility differentials by educational status, as one's schooling is often prior to one's childbearing. Though observed as a block at a point in time, these variables can nevertheless be causally ordered on the basis of one's knowledge of the field. Appealing to all available knowledge is essential.

When the variables cannot be causally ordered, due to insufficient background knowledge or to a lack of information on the temporal sequence of events, an exploratory analysis of the data and especially of the so-called *Big Data* (in the sense of very large structured and unstructured data sets) can possibly be helpful in revealing changing characteristics over time and suggesting the temporal sequence of events. However, exploratory analysis, such as data mining or machine-learning of big data bases, yields solely (possibly spurious) associations among variables and not causal links. An exploratory data approach is never a substitute for sound causal modelling, such as the framework presented in this paper, but it can usefully inform it, especially when background knowledge on the topic of interest is scant.

5 Concluding remarks

Causal methods are an essential part of the scientific enterprise. As Illari and Russo (2014) notice, they are many, tailored to different scientific contexts, and to different concepts of causality. The social sciences are also confronted with such diversity. In this paper, we offer a synthesis of an approach developed over the last decade or so by the authors. The overall goal is to provide an account of how, in a number of social science disciplines—and especially those using observational data—social scientists can go beyond mere description, based essentially on statistical associations, and propose a causal explanation of a phenomenon of interest.

We aim to give a broad methodological framework that reconstructs the whole modelling process and identifies key moments. Our approach pays attention to particular statistical features of structural causal models, for instance to the role of stochasticity in representing an actual DGP, to distributions, or to the identification of sub-mechanisms with conditional distributions. With the proposed framework, statistical modelling is part of what is at stake, but does not exhaust the meaning of structural causal modelling. Adopting an SCM approach means to endorse a particular view on modelling in general (the hypothetico-deductive methodology), and a specific stance on exogeneity (namely as a condition of separability of inference) and to interpret marginal–conditional decompositions (namely as mechanisms).

A fundamental role is given to background knowledge as preliminary information for the purpose of elaborating the recursive decomposition and allowing, in particular, to interpret each conditional distribution as a sub-mechanism of the DGP. Moreover, a subsequent analysis of structural stability leads to identify possible heterogeneity in the population of reference.

One further question is whether such a model merely represents causal relations out there, or whether it contributes to *building* causal knowledge. It should be clear from the previous discussion that we take structural models to be important *tools* to build knowledge of social phenomena. Our concept of causality is thus largely epistemic, in

the sense that it is related to the way we get to know and understand social reality. We don't deny that causal relations happen in the world—a kind of metaphysical stance about causality—but we think that *if* we can get any grip on the 'true' causal relations in the world, this is thanks to our modelling enterprise, which sometimes succeeds and sometimes fails. In this sense, the epistemological foundations that we offer, and in particular the role of variation, are essential in that it tells us *how* to search for causes: by looking for and testing for suitable variations in the data sets. But variations are, alone, not enough. This is why another important element in our approach is the interpretation of the marginal–conditional decomposition as a mechanism: it contributes to *explain* the phenomenon of interest. The structural modelling framework hereby proposed is thus well in line with the most recent developments in the philosophy of science, in which mechanisms play a more prominent role than laws, at least for explanation.

In sum, notwithstanding the major advancements in quantitative methods in social science, the structural modelling framework that we propose aims to offer broad methodological and epistemological tools to researchers approaching various questions in social research. Establishing causal relations and offering causal explanation of social phenomena is not just a matter of mastering the technicalities of sophisticated statistical models, but also—or perhaps most importantly—of having a solid methodological approach, with which to make relevant choices at each stage of the modelling process. In our intention, structural causal modelling can offer that.

Acknowledgements Comments, in particular by Catherine Gourbin, Renzo Orsi, and Frans Willekens, on former versions of this paper, are gratefully acknowledged.

References

Fagiolo, G., Moneta, A., Windrum, P.: A critical guide to empirical validation of agent-based models in economics: methodologies, procedures, and open problems. Comput. Econ. **30**, 195–226 (2007)

Gourbin, C., Wunsch, G., Moreau, L., Guillaume, A.: Direct and indirect paths leading to contraceptive use in urban Africa. An application to Burkina Faso, Ghana, Morocco and Senegal. Rev. Quetelet/Quetelet J. **5**(1), 33–70 (2017)

Hood, W.C., Koopmans, T.C. (eds.): Studies in econometric methods, Cowles Foundation Monograph 14. Wiley, New-York (1953)

Illari, P., Russo, F.: Causality: philosophical theory meets scientific practice. Oxford University Press, Oxford (2014)

Johnson, R.B., Russo, F., Schoonenboom, J.: Causation in mixed methods research: the meeting of philosophy, science, and practice. J. Mixed Methods Res. (2017). https://doi.org/10.1177/1558689817719610

Koopmans, T.C.: Measurement without theory. Rev. Econ. Stat. **29**, 161–173 (1947)

Koopmans, T.C. (ed.): Statistical inference in dynamic economic models, Cowles Foundation Monograph 10. Wiley, New York (1950)

Little, D.: Levels of the social. In: Risjord, M., Turner, S. (eds.) Philosophy of anthropology and sociology, pp. 343–371. Elsevier Science, Amsterdam (2006)

Mackie, J.L.: Causes and conditions. Am. Philos. Q. **2**(4), 245–264 (1965)

Morgan, S.L., Winship, C.: Counterfactuals and causal inference. Cambridge University Press, Cambridge (2007)

Mouchart, M., Orsi, R.: Building a structural model: parameterization and structurality. Econometrics **4**, 23 (2016). https://doi.org/10.3390/econometrics4020023

Mouchart, M., Russo, F.: Causal explanation: recursive decompositions and mechanisms, chap. 15. In: McKay Illari, P., Russo, F., Williamson, J. (eds.) Causality in the sciences, pp. 317–337. Oxford University Press, Oxford (2011)

Mouchart, M., Russo, F., Wunsch, G.: Structural modelling, exogeneity, and causality, Chapter 4. In: Engelhardt, H., Kohler, H.-P., Fürnkranz-Prskawetz, A. (eds.) Causal analysis in population studies: concepts, methods, applications, pp. 59–82. Springer, Dordrecht (2009)

Mouchart, M., Russo, F., Wunsch, G.: Inferring causal relations by modelling structures. Statistica **LXX**(4), 411–432 (2010)
Mouchart, M., Wunsch, G., Russo, F.: Controlling variables in social systems: a structural modelling approach. Bull. Sociol. Methodol. **132**, 5–25 (2016)
Pearl, J.: Causality. Models, reasoning, and inference. Cambridge University Press, Cambridge (2000). **(revised and enlarged in 2009)**
Russo, F.: Causality and causal modelling in the social sciences: measuring variations, Methodos Series, vol. 5. Springer, Berlin (2009)
Russo, F., Wunsch, G., Mouchart, M.: Inferring causality through counterfactuals in observational studies. Some epistemological issues. Bull. Sociol. Methodol. **111**, 43–64 (2011)
Wold, H.O.: Causality and econometrics. Econometrica **22**(2), 162–177 (1954)
Wunsch, G., Mouchart, M., Russo, F.: Functions and mechanisms in structural-modelling explanations. J. Gen. Philos. Sci. **45**(1), 187–208 (2014)
Wunsch, G., Mouchart, M., Russo, F.: Causal attribution in block-recursive systems: a structural modelling perspective. Methodol. Innov. (2018). https://doi.org/10.1177/2059799118768415
Wunsch, G., Russo, F., Mouchart, M.: Do we necessarily need longitudinal data to infer causal relations? Bull. Sociol. Methodol. **106**, 5–18 (2010)