J Gen Philos Sci (2014) 45:187–208
DOI 10.1007/s10838-014-9242-z

ARTICLE

Functions and Mechanisms in Structural-Modelling Explanations

Guillaume Wunsch · Michel Mouchart · Federica Russo

Published online: 9 February 2014

Abstract One way social scientists explain phenomena is by building structural models. These models are explanatory insofar as they manage to perform a recursive decomposition on an initial multivariate probability distribution, which can be interpreted as a mechanism. Explanations in social sciences share important aspects that have been highlighted in the mechanisms literature. Notably, spelling out the functioning the mechanism gives it explanatory power. Thus social scientists should choose the variables to include in the model on the basis of their function in the mechanism. This paper examines the notion of 'function' within structural modelling. We argue that 'functions' ought to be understood as the theoretical underpinnings of the causes, namely as the role that causes play in the functioning of the mechanism.

Keywords Causality · Explanation · Function · Mechanism · Recursive decomposition · Structural modelling

1 Introduction

Social scientists study a variety of phenomena both at the group and at the individual level. One difficulty social scientists face, however, is that their object of investigations—human

G. Wunsch
Demography, Université catholique de Louvain, Louvain-la-Neuve, Belgium
e-mail: g.wunsch@uclouvain.be

M. Mouchart
CORE & ISBA, Université catholique de Louvain, Louvain-la-Neuve, Belgium
e-mail: michel.mouchart@uclouvain.be

F. Russo (✉)
Dipartimento di studi umanistici, Università di Ferrara, Ferrara, Italy
e-mail: federica.russo@unife.it

beings and their behaviour—can be very different across time, across cultures, across geographic regions, or, of course, across individuals. Consider for instance alcohol consumption. Balt people drank, in the former USSR, because drinking relieved stress. British people drink to socialise. Mediterraneans drink to accompany a lavish meal, although binge drinking is becoming more and more popular among young people. Alcohol consumption, it seems, has different functions in different places. Drinking has, ceteris paribus, the same health effects, but also different 'social' reasons and effects. We can have a better understanding of drinking behaviours if we find out the function of drinking in the specific reference class under analysis. Social scientists seek to explain phenomena such as these ones, studying the differences and similarities across populations. However, in social research we encounter various traditions to tackle explanation.

To illustrate a widespread approach to explanation in the social sciences, we have examined all articles published by a well-known and respected population journal, the *European Journal of Population*, during the past twelve months at time of writing, i.e. the four issues of 2012. Broadly speaking, in most cases the authors include, based on a literature review, all known and observable determinants X_i of an outcome variable Y into a single-equation model, and then consider the impact of each of these variables on Y independently of the other predictor variables, sometimes using a stepwise approach. A distinction is often made between key explanatory variables and controls (though not usually specifying if the latter are confounders, mediators, moderators, or independent covariates), less often between individual and contextual variables, and a theoretical framework is frequently developed in order to present the main research questions and hypotheses, though a full explanatory mechanism is most often absent. No causal ordering of the predictor variables is generally attempted. All predictor variables are implicitly considered in this approach as if they were independent from one another, though some authors do use some form of classification analysis to examine possible groupings among variables, and interactions between variables are sometimes examined. This view is nevertheless largely based on associations informed by background knowledge, and few authors attempt to hierarchize their explanatory variables by way of a causal ordering.

There are of course exceptions to this standard approach, and more exceptions would be found if the set of texts were enlarged and additional journals considered. Indeed, there is also a tradition in the social sciences where the researchers take pain to specify the network, or system, of determinants and effects in a causal approach. This is the 'structural modelling' approach that we will consider in this paper. Yet, the *reasons* for including some variables rather than others in the system are often not sufficiently justified and supported by argument. Furthermore, the explanatory relevance of the network of variables itself is usually lacking. For instance, variables are many a time poorly conceptually defined, and the putative causal relations among variables often lack strong evidential support. Consequently, there is a pressing need for improving the scientific practice by specifying the *mechanism* underlying the observed phenomena and the *functions* that the variables have in such a system. This is where we start building a bridge between the current scientific practice and the existing philosophical literature on mechanisms.

In fact, according to an established tradition in philosophy, to explain a phenomenon we need to invoke the causes that bring it about. Though other forms of explanation are possible, such as systemic or evolutionary explanations, in this paper we consider causal explanation. Recent debates have shown that, arguably, in order to explain a phenomenon we need more than to invoke its causes. We need to provide a (plausible) *mechanism* for the phenomenon. This idea has been defended both in the literature examining mechanisms in biology and neuroscience (see, among others, the works of Bechtel and Abrahansen 2005, Craver 2007, Glennan 1996, 2002, etc. discussed later in the paper) and in the literature on mechanisms in the social sciences (see for instance Hedström and Swedberg 1998; Demeulenaere 2011a; or Little 2011).

The 'structural models' we will consider in this paper belong to this tradition. However, the *conceptual* links between this modelling framework and mechanistic explanation have not been thoroughly examined yet. Social scientists use structural models to explain phenomena and these models are explanatory insofar as they model mechanisms. In this paper, following Mouchart et al. (2010), we argue that, from a methodological point of view, the bulk of a structural-modelling *explanation* is the concept of recursive decomposition. This means breaking down a global mechanism, involving all variables of interest altogether, into an ordered sequence of simpler sub-mechanisms, each one involving an endogenous variable and exogenous variables, such that exogenous variables can be explicitly interpreted as causes of the endogenous one (see Sect. 2.2.2). Additionally, this approach represents mechanisms through (conditional) distributions rather than through equations, accordingly providing more flexibility and more subtleties in the evaluation of the effects of a cause on an outcome variable.

However, in order to provide a sound explanation researchers need to go a step further. Besides modelling the mechanism and identifying the relevant causes acting in it, they also have to say *why* some (putative causal) variables have been chosen instead of some others. We will argue that this amounts to spelling out the *function*, or role, these variables play in the mechanism. This paper examines the meaning and use of 'function' in structural-modelling explanations.

The paper is organised as follows. Section 2 presents structural-modelling explanations, with particular attention to the recursive decomposition and to the overall model building methodology. Section 3 considers the relation between mechanisms and explanation and prepares the ground for the discussion of the concept of function in Sect. 4. In particular, we endorse an epistemic view of explanation, according to which it is the *description* (of the functioning) of the mechanism that does the explaining job. We also explain why the recursive decomposition may be interpreted as a (social) mechanism. Finally, in Sect. 4 we examine two accounts of 'function': in terms of 'role-function' and 'isolated descriptions' on the one hand, and in terms of 'functional architecture' on the other hand. We argue that the appropriate concept is that of 'role-function', and that modelling a mechanism requires individuating both the function of the mechanism and of its components. This is done throughout the modelling procedure by providing the theoretical underpinnings, i.e. the functions of the chosen causal factors.

2 Structural-Modelling Explanations

Structural models, in a broad sense, are widely used in current social science practice. The details of the *meaning* of 'structural' or of the specific technical characteristics may vary among different schools of thought and authors. In this section we trace the main lines of the origins of structural models, in order to introduce the approach that we endorse in this paper. Such an overview will also set the ground for presenting the conceptual links between structural models and mechanistic explanation through the recursive decomposition.

2.1 Origins of Structural Models

In econometrics, the members of the Cowles Commission, in particular Koopmans (1950a) and Hood and Koopmans (1953), developed a concept of structural model under the motto 'no measurement without theory' (Koopmans 1947). They proposed to start with a set of equations generated by economic theory, possibly enlarged in order to incorporate heterogeneity among the agents. This background implies which variables should be considered as globally exogenous and the other ones being considered endogenous. These equations form a system of simultaneous equations qualified as 'structural', with the aim of representing an economic structure. In most cases, simultaneity among equations is rooted in mechanisms of equilibrium linking the behavior of particular agents modeled by each equation. From this structural model they deduced the joint distribution of the endogenous variables conditionally on the exogenous variables and called this model the 'reduced form' (Koopmans 1950b). They analysed a substantial issue of identification of the structural parametrization and pointed out that economic theory should provide restrictions necessary to ensure the identification of the structural parameters (Koopmans 1953). Economic theory also imposes restrictions (named 'over-identification') on the parameters of the reduced form. Briefly, in this approach economic theory is the basis for constructing and identifying the structural model and is also used to improve statistical inference thanks to over-identifying restrictions.

Among many contributions, Herman Wold (1949, 1954) followed up the Cowles Commission works on simultaneous equations in the direction of a recursive decomposition; his purpose was to disentangle the simultaneity of a system of structural equations. He argued that economic theory should provide a naturally recursive system but that simultaneity may derive from incomplete observability of the behaviour of the system. Further developments in simultaneous equation models led Wold and his collaborators to develop a class of models close to the structural models of other social sciences. The role of recursivity is not only to disentangle simultaneity, but it is also at the root of the concept of explanation, in the sense that explaining a complex mechanism is based on a decomposition into simpler sub-mechanisms. This is the view that we will advance in Sect. 2.2 and that will set the ground for discussing the meaning of functions in Sect. 4.

Structural models also originate in the path analytic methodology developed by Sewall Wright and in the subsequent causal models for non-experimental research—from the 1960s and 1970s—of Hubert Blalock, Otis Dudley Duncan, and Raymond Boudon, among others. In a series of papers published from 1918 onwards, Sewall Wright developed the methodology of *path analysis* for measuring causal relations. He represented these relations in branching sequential order (i.e. as a recursive system) by an arrow-diagram, a graphic representation taken up much later and expanded by Judea Pearl, especially in his directed acyclic graphs (DAG) approach to causality (Pearl 2000). Wright (1934, 175) took pains in stressing the fact that his path analytic model should be based on "such information as may be at hand with regard to causal relations", thus making the model dependent upon background knowledge, unlike data mining or exploratory data analysis.

The issue that the starting point of causal modelling is "a firm grasp of the relevant scientific theory", and not the statistical methods actually employed, is also highlighted by Otis Dudley Duncan (1975, 6) in his application of structural equation modelling to sociological data. Hubert Blalock too was influential in linking theory and empirical research in the social sciences by the use of causal models, though he insisted that

> […]we can never actually demonstrate causal laws empirically. This is true even where experimentation is possible. Causal laws are working assumptions of the scientist, involving hypothetical statements of the *if-then* variety. Included among the *if's* of causal assertions is the supposition that all relevant variables have been controlled or can safely be ignored. This kind of assumption can never be tested empirically (Blalock 1964, 172–173).

Raymond Boudon (1967) took the same stance that statistics alone cannot lead to establishing causal relations in the absence of a causal interpretation. Causal claims always relate to the causal model that is proposed. His "dependency analysis" of causal structures was influential in introducing the American structural modelling approach and causal diagrams (actually DAGs) to the French-speaking social scientists.

2.2 From Structural Models to Explanation

The structural modelling approach adopted in this section is a heir of the tradition just presented, and it also goes beyond, in that it puts the recursive decomposition at the core of explanation. The recursive decomposition allows us to distinguish different causal roles for the variables, and for this reason—together with the concept of function developed later in Sect. 4—it is crucial for understanding the explanatory import of a mechanism.

2.2.1 Structural Models: A Further Development

As just mentioned, an important development in structural modelling is to provide a thorough understanding of the recursive decomposition and of its explanatory import. In this section, we present an account of model-building, that is *how* a structural model has to be built in order to explain phenomena. To do so, the researcher builds a recursive decomposition, as illustrated next in Sect. 2.2.2. In particular, we stress the *explanatory* role of the recursive decomposition: it has to be interpretable as a mechanism (see also Mouchart et al. 2010). This feature will be necessary to understand the meaning of functions, as discussed later in Sect. 4.

We follow in the footsteps of the father-figures of social science methodology (e.g. Blalock), departing from the Cowles Commission tradition, in that we enlarge the role of theory to the wider concept of background knowledge as follows. First, the structural modeller does not restrict background knowledge to economic theory or to any other particular theory, but s/he takes into account all relevant information; second, the structural modeller navigates among plausible specifications of the model, guided by invariance considerations across specified changes of the environment. As a consequence, building a structural model is a more complex issue than just a statistical translation of a pre-specified theory.

However, in contradistinction to the models based on a system of (linear) *equations* that are at the basis of most of the works of the father-figures of social science methodology, we develop a general presentation of structural modelling in terms of conditional *distributions* and recursive decomposition of the data set. Switching from equations to distributions is important for the following reason. Writing an equation endowed with a residual term does not have statistical meaning, unless it is completed with additional specifications providing a precise interpretation in terms of the implied conditional distributions. For instance, in the literature on econometric simultaneous equation systems, it has been crucial to understand that a structural equation does not represent a conditional expectation (for a recent discussion, see e.g. Fennell 2011). Moreover, according to the mechanism, a cause may have a more substantial effect on the conditional variance or on the conditional tails than on the conditional expectation (see e.g. Heckman 2008).

We restrict the present discussion to recursive systems without reciprocal or two-way causation. As other social scientists have done (see e.g. Morgan and Winship 2007, chapter 8), we defend the view that causal explanations require modelling a plausible mechanism for the data generating process. In this perspective, the condensed recursive decomposition can be interpreted as a mechanism (see later Sect. 2.2.2). However, to support such an interpretation, we also require that the whole mechanism, and each of its components (or sub-mechanisms), have a specified function, as discussed later in Sect. 4. Furthermore, we stress that the characteristics of the structural model should be stable in a given context, in order to reach explanatory generalisations. Finally, a structural modelling approach is not necessarily rooted in a counterfactual/manipulationist basis, though it is not in contradiction with the latter: counterfactuals and interventions can indeed be useful tools for assessing causal relations (see for instance Russo et al. 2011). The issue, in fact, is not to replace manipulations with something else, but to offer an account that is broader in scope. It is worth noting, in particular, that counterfactuals are not based on explicit mechanisms while this is what structural modelling does.

We also develop a specific meaning of 'structural' that departs from other scholars' use in the current debates. Our meaning of 'structural' allows us to 'graft' the concept of function to it. Instead, the accounts proposed by e.g. Woodward (2003, ch. 7), Pearl (2011), or Hoover (2011) do not take issue *specifically* with what 'structural' means and do not spell out the conceptual underpinnings of the recursive decomposition in terms of mechanistic explanation. The contributions of the above-mentioned authors confine the discussion of their concept of structural modelling (1) to the relation with the notion of invariance under intervention (Woodward), or (2) to a general framework supposedly able to subsume, as special cases, particular models such as structural equation models, potential outcome models, and graphical models (Pearl), or (3) to relations with the notion of causal ordering developed by Herbert Simon (Hoover). For a discussion, see Russo (2011a).

2.2.2 Structural Models: Explaining Through the Recursive Decomposition

In the context of quantitative analysis, our perspective starts from statistical models typically made of a parametrized family of distributions generated by a set of plausible hypotheses; these suggest how the observations under analysis may have been produced by the so-called data generating process (DGP).

Modelling the specific pattern underlying the data generating process requires taking into account the contextual knowledge of the field of application. A structural model should accordingly help understand the data generating process; more precisely, its characteristics, or parameters, should be interpretable. A structural model should also be stable, or invariant, across changes of the environment (we use 'invariance' and 'stability' interchangeably).

The complexity of the social world typically leads to a multidimensional approach. Understanding such complex phenomena requires introducing structure into the model and this can be achieved by operating an ordered decomposition of the multivariate distributions of the variables. In other words, explaining involves breaking down a complex phenomenon into simpler parts. This decomposition of the data generating process is operationalised by a systematic *marginal-conditional decomposition* as follows.

Let us introduce the approach through a simple example. Suppose we examine the impact of cigarette smoking on the incidence of cancer of the respiratory system. Background knowledge tells us that we must take into account the fact that exposure to asbestos dust is also a cause of lung cancer. The roles of smoking and of asbestos exposure in the development of lung cancer are now well-known and are due to biological mechanisms (Biello et al. 2002). Tobacco smoke contains more than 60 different toxic substances, which can cause the development of cancer. As to asbestos exposure, it may lead to malignant mesothelioma, a cancer of the mesothelial cells in the lining covering the lungs.

Exposure to asbestos dust and smoking are associated, i.e. proportionally there are more persons exposed to asbestos in the smoking group than in the non-smoking group. But why are smoking and asbestos exposure associated? Studies in demography and in epidemiology have shown that both smoking and asbestos exposure are dependent upon one's socio-economic status (*SES*): those with a lower *SES* tend to smoke more and to work more often in unhealthy environments than those with a higher *SES*. In fact, people with a low SES are more exposed to asbestos fibres in the workplace than individuals with a high SES, as they are more prone to manipulating products containing asbestos. Explanations of the socioeconomic differences in smoking include lack of knowledge, scarce material resources, psycho-social stress due to an unfavourable social position and poor material conditions (Laaksonen et al. 2005). So *SES* is a key causal variable in the model for the *practices*, i.e. the social mechanisms, that it induces.

This simplified example is discussed in Mouchart et al. (2009) but a real study would also consider other causal factors and paths, and the interaction between smoking and asbestos exposure (Saracci 2006). A real analysis might also take into account intermediary mechanisms (mediators) between each of the explanatory variables. In this paper, we set aside the problem of how far or deep a researcher has to go in order to provide an explanation of a given phenomenon; this also depends on available information and on the relevance of the mediators for the chosen research question. For example, one could specify a mechanism about economic behaviour explaining the direct arrow between SES and T in Fig. 1: the choice of cigarette (e.g. with or without filter) also depends on how much the person can afford to spend. But the specification of such mechanisms may not be possible (lack of data) or relevant for demographic purposes, though it may be relevant for other reasons (economic or policy analysis).

A *directed acyclic graph* (DAG) can be drawn as in Fig. 1, where A represents exposure to asbestos, T tabacism, and C cancer incidence. Figure 1 incorporates two assumptions grounded on background knowledge, namely: $A \perp\!\!\!\perp T \mid SES$ and $C \perp\!\!\!\perp SES \mid A, T$, meaning that the distribution of T given SES and A depends only on SES (or, equivalently the distribution of A given SES and T depends only on SES) and that the distribution C given A, T and SES depends only on A and T.

The first hypothesis is depicted by the absence of arrow between A and T, although tabacism and asbestos exposure are in fact not independent from one another as they are both related to one's SES, i.e. they have a common cause. The second hypothesis leads to an absence of an arrow between SES and C: the impact of SES on C is mediated by A and T.

It should be noted that in this case background knowledge leads to ordering the 4 variables in such a way that the recursive (marginal-conditional) decomposition of their joint distribution, namely:

$$p_{SES,A,T,C} = p_{SES}\, p_{A|SES}\, p_{T|SES}\, p_{C|A,T} \tag{1}$$

is such that each factor represents a contextually meaningful sub-mechanism. More

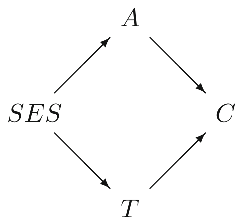

Fig. 1 Socio-economic status, smoking, asbestos exposure and cancer of the respiratory system

specifically, $p_{C|A,T}$ represents the biological mechanism whereas the product $p_{A|SES}\,p_{T|SES}$ represents the social mechanism.

More generally, let us consider a decomposition of data set X into p components such as $X = (X_1, X_2, \ldots X_p)$. Suppose that the components of X have been ordered in such a way that in the complete marginal-conditional decomposition

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

where $\theta_{j|1,\ldots j-1}$ stands for the parameter characterising the conditional distribution of $(X_j \mid X_1, X_2, \ldots X_{j-1})$, the factors of the right hand side of (2) have mutually independent parameters, i.e. in a sampling theory framework:

$$\omega = (\theta_{p|1,\ldots p-1}, \theta_{p-1|1,\ldots p-2} \cdots, \theta_1) \in \Theta_{p|1,\ldots p-1} \times \Theta_{p-1|1,\ldots p-2} \cdots \times \Theta_1. \tag{3}$$

Condition (3) excludes restrictions among parameters of different factors of the right hand side; in particular common parameters of different factors are not allowed. Under the condition (3), the decomposition (2) is called a *recursive decomposition*.

Once the number p of components increases, background knowledge, substantiated by analysis of the data and statistical tests, can provide a simplification of the factors in the form of conditional independence properties, as has been done in the above example on lung cancer. More specifically, it is typically the case that the distribution of $(X_j \mid X_1, \ldots, X_{j-1})$ is known not to depend on some of the conditioning variables. Thus there is a subset $\mathcal{I}_j \subset \{X_1, \ldots, X_{j-1}\}$ of variables whose actual relevance for the conditional process generating $X_j \mid X_1, \ldots, X_{j-1}$ is defined by the property

$$X_j \perp\!\!\!\perp X_1, \ldots, X_{j-1} \mid \mathcal{I}_j, \theta. \tag{4}$$

This property implies that the factor $p_{X_j|X_1,X_2,\ldots X_{j-1}}$ in (2) is actually simplified into $p_{X_j|\mathcal{I}_j}$ and $\mathcal{I}_j$ may be called the *relevant information of the j-th factor*. Once $\mathcal{I}_j$ has been specified for each factor, (2) is condensed into

$$p_{X_1,X_2,\ldots X_p|\theta} = \prod_{1 \le j \le p} p_{X_j|\mathcal{I}_j,\theta_{j|1,\ldots j-1}} \tag{5}$$

This form will be called a *condensed recursive decomposition*. Thus Eq. (1) in the example is a condensed recursive decomposition because it is in the form of (5) rather than in the form of (2).

Recursive decompositions raise two important issues. The first is whether the endogenous variable of each factor in Eq. (2) is univariate or not. If each X_j is univariate, the decomposition is said *complete*; whereas if some X_j is a vector of variables, the decomposition is said *partial*, or block-recursive. This means that for X_j the modelling has not succeeded in decomposing the generation of X_j into sub-mechanisms generating each component of X_j. In a DAG perspective, complete or partial refer to the question of whether each node corresponds to a univariate random variable or to a vector of random variables.

The second issue is whether the interpretation of the factors of the recursive decomposition in terms of sub-mechanisms requires that the conditioning variables of X_j are all the "ancestors", i.e. $X_1, X_2, \ldots, X_{j-1}$ as in Eq. (2), in which case the decomposition is said *saturated*. Conversely, if some of the ancestors are absent, as in Eq. (5), the decomposition is said to be *condensed*. In a DAG perspective, a saturated decomposition is represented by a graph with all the possible arrows, provided that the "directed" and "acyclic" properties are respected. A condensed decomposition is represented by a graph where some of the potential arrows have been deleted on the ground of background information leading to assumptions about conditional independence, as in Eq. (4).

The goal of structural modelling is to characterise, as much as possible, clearly identified and interpretable sub-mechanisms of the global mechanism. More specifically, Eqs. (5) and (3) represent a global mechanism; its factors, given by the conditional distributions, are to be interpreted as sub-mechanisms. The choice of the marginal-conditional decomposition is therefore not arbitrary; we need background knowledge and invariance to select one among the p! possible decompositions. In other words, the marginal-conditional decomposition *alone* does not provide a (causal) explanation of a given phenomenon, but the whole modelling procedure does (Mouchart and Russo 2011). In Sect. 4, we discuss what it means that variables have a function in these mechanisms and sub-mechanisms. This section and the following one focus instead on the interpretation of the recursive decomposition as a mechanism, and on its explanatory import (Russo 2011b).

A crucial aspect of the multiplicative structure (5) under the condition (3) is to allow an inference on the parameters of interest, namely functions of $\theta_{j|1,\ldots j-1}$, *independently* of the specification of the process that generates the conditioning variables, provided only that condition (3) is satisfied. This is accordingly a condition of separation of inferences. In other words, the parameters of interest should only depend upon the parameters identified by the conditional model and the parameters identified by the marginal process should be "independent" of the parameters identified by the conditional process. In this case, the conditioning variables of each conditional component of the decomposition are *exogenous* variables (for details, see Koopmans 1950b, Florens, Mouchart and Rolin 1980, Engle, Hendry and Richard 1983, Florens and Mouchart 1985).

Mouchart et al. (2009) argue that *causes* may then be viewed as exogenous variables in the condensed recursive decomposition, alternatively as the relevant information of a structurally valid conditional distribution. Thus the variation-free condition (3) does not only allow us to separate the inferences on $\theta_{j|1,\ldots,j-1}$ and on $\theta_{1,\ldots,j-1}$, but it also allows us to distinguish the process generating the causes, characterised by $\theta_{1,\ldots,j-1}$, and the process generating the effects, characterised by $\theta_{j|1,\ldots,j-1}$. Separating causes from effects mirrors the asymmetry of causation. Associating each component of the condensed recursive decomposition with a sub-mechanism of the data generating process provides interpretability to the parameters of the corresponding conditional distributions. Readers familiar with the literature on graph models may recognise that a directed acyclic graph (DAG) is a simplified graphic representation of a condensed recursive decomposition and that the causal structure is depicted by the set of ancestors. This simplification may not be suitable to complex situations.

Why interpreting *exogenous* variables as *causal* factors? The reason is that, as we discussed earlier, structural modelling aims to explain a complex phenomenon. In turn, to explain a complex phenomenon, we need to model 'structures', namely mechanisms where the acting parts are interpreted as causes. The philosophical underpinnings of such an approach are the following:

1. Explanation, at least in the social contexts we are concerned with, is *mechanistic* explanation. (Notice that mechanisms need not be deterministic. In a structural modelling framework, mechanisms are, rather, stochastic.)
2. Mechanistic explanation is successful insofar as the *functioning* of the mechanism is properly spelled out.
3. Spelling out the functioning of the mechanism amounts to identifying the *causes*, their actions, and their effects, and to provide a *justification* for the inclusion of some specific variables, rather than others, in the recursive decomposition.

In Sects. 3 and 4 we spell out these philosophical underpinnings. In particular, Sect. 3 focuses on the relation between structural models and explanation, and Sect. 4 on the meaning of function.

3 Explanation and Mechanisms

We mentioned earlier that a structural model explains insofar as it models a (plausible) mechanism for the phenomenon of interest. But how is this relation between explanation and mechanisms to be understood? In this section we give some background of how this issue has been addressed in the literature and of how it is addressed in the context of structural explanations.

3.1 Why Mechanisms Explain

The recent and rapidly expanding literature on mechanisms is working simultaneously on two (interrelated) fronts. On the one hand, mechanists (i.e. scholars working in the philosophy of mechanisms) work towards a definition of what a mechanism is. On the other hand, they also work towards spelling out the explanatory import of mechanisms. These two issues are of course related.

Several definitions of 'mechanism' have been proposed in the literature (Machamer et al. 2000; Glennan 2002; Bechtel and Abrahamsen 2005; Franck 2002). We embrace the one given by Illari and Williamson (2012):

> A mechanism for a phenomenon is composed of entities and activities organized so that they are responsible for the phenomenon

Illari and Williamson offer several arguments to say that their definition constitutes a potential 'consensus'. We lack space to thoroughly discuss why we agree with Illari and Williamson that it represents a consensus definition. It will suffice to say that this definition is general enough to be applied to social contexts, besides biology or neuroscience (the main subject fields in which the mechanism debate takes place). It is also worth mentioning that Illari and Williamson's definition still hinges (as indeed the other definitions did) upon the 'entity-activity' ontology and stresses the importance of the organisation of the mechanisms. We shall see later in Sect. 4 why these components (entity-activity and organisation) are important in order to understand the concept of 'function' in structural explanations. Another reason for our choice is that Illari and Williamson do not define mechanisms in terms of manipulation or counterfactuals, and this makes it particularly suited to the structural modelling framework developed in the previous section, which is also not based on manipulations or counterfactuals.

It is worth noting that the philosophical literature on mechanisms originated from (and to some extent still draws on) paradigmatic examples of mechanisms in biology or neuroscience. Comparatively less attention has been paid to social mechanisms. Thus, by extending the arguments developed for biological or psychological mechanisms, we also endeavour to bring social mechanisms into this philosophical debate.

A detailed discussion of social mechanisms, from the viewpoint of analytical sociology especially, can be found in a volume edited a few years ago by Hedström and Swedberg (1998), on which the following is partly based. Little (1991, 2011) also offers a discussion of social mechanisms drawing on examples from various social science disciplines. As mentioned in the definition above, generally speaking, a mechanism is conceived of a set of interacting parts or elements organised in such a way that it produces or generates an effect. Following this line of reasoning, a social mechanism is then a plausible account of how change in social processes is brought about through sequences of cause and effect relations. We look for mechanisms in order to increase our understanding of the observed joint variations between variables of interest and to possibly intervene with policy actions at the right point of the mechanism (Russo 2009).

Many social mechanisms are rooted in the principle of methodological individualism, i.e. in the behaviours of individuals. They can relate to the interactions among individuals (micro-micro effects), to the impact of social institutions and contexts on individual behaviours (macro-micro effects), or to the consequences of individual actions on the macro level (micro-macro effects). From this viewpoint, all macro-macro effects should be explained by the actions of individuals: there would be no such thing as macro-level mechanisms (for an opposite view, see however Ylikoski 2011). Social mechanisms are of course unobserved analytical constructs that propose hypothetical causal links between observed events. As stressed e.g. by Diego Gambetta (1998), we hypothesise social mechanisms but we cannot observe them directly. This is indeed what the recursive decomposition discussed in Sect. 2.2.2 is supposed to do: to model data generating processes in terms of mechanisms. We can however test mechanisms via their predicted effects, i.e. by comparing implications of the mechanisms with empirical facts.

It is worth noting that we are not just saying that there *is* a data generating process or data generating mechanism that 'produces' the data. *Modelling* the data generating process precisely means to specify what this mechanism is. Part of the modelling procedure also concerns choosing the variables according to the function they have—Sect. 4 provides a thorough discussion about this claim. The causal structure modelled in the recursive decomposition is meant to open the black box—to the best of our knowledge—namely it aims to explain associations between variables by plausible cause and effect relations generating these associations.

Consider the following example. An association between two variables, say occupation and fertility, could be explained by invoking a mechanism from occupation to fertility, but also invoking a 'reverse' mechanism from fertility to occupation, or even invoking an altogether different mechanism where another variable, say education, has both occupation and fertility as effects. A social explanation of the observed association between occupation and fertility should indeed specify what is the mechanism behind. But it may be the case that in *different* populations *different* mechanisms are at work.

So a social mechanism is an explanation of how change in an outcome variable, or a set of outcome variables, is brought about, and it also requires specifying the *functioning* of the mechanism. Such an explanation should display a suitable level of stability in order to reach some level of explanatory generalisation.

From the point of view of structural-modelling explanations there are two aspects that need to be discussed. The first is modelling social mechanisms by means of the recursive decomposition, which has been detailed earlier in Sect. 2.2.2. The second is the function that the mechanism and that its elements play in the explanation. This is undertaken later in Sect. 4.

3.2 Social Mechanisms as Recursive Decompositions

The bulk of a structural explanation is the recursive decomposition. This amounts to breaking down an initial multivariate probability distribution into 'smaller pieces'. We say that the whole recursive decomposition represents a global mechanism, whereas the smaller pieces represent sub-mechanisms.

For instance, Gaumé and Wunsch (2010) carry out a study on the determinants of subjective health in the Baltic States for the period 1994–1999 and offer a mechanism, based on background knowledge, that explains subjective health by a series of possible individual determinants and their interrelations. The global mechanism explaining subjective health has sub-mechanisms. For instance, the impact of alcohol consumption on subjective health depends upon educational level of the individual, physical health status, level of psychological distress. These items themselves depend upon other factors that are part of the global mechanism; for example, psychological distress is influenced by the individual's locus of control and social support. A number of remarks are in order.

First, *different* mechanisms may explain a same phenomenon. In the Baltic study just mentioned, researchers attribute to alcohol consumption a function of 'stress-relieving' in the mechanism explaining subjective health outcomes. But alcohol consumption can have different functions in different contexts, as a report of The Social Issue Research Centre (1998) shows.

Second, not all mechanisms can be represented by a recursive decomposition. For instance, we doubt that the mechanism for photosynthesis can be represented in terms of a recursive decomposition. We do not claim either that all social mechanisms can be represented by recursive decomposition, due to a contingent lack of information. For example, Mouchart and Vandresse (2007) analyse data on a set of commercial contracts, each one characterised by a price and some characteristics of the contract and claim that knowing only the contract resulting from negotiations does not provide sufficient information for deciding which of the characteristics or the price were firstly decided upon, the most likely hypothesis being that each contract was concluded after several rounds of negotiations. The presence of latent (i.e. unobserved) variables may also play havoc with the results of the structural model.

The researcher endeavours to perform a *complete* recursive decomposition (in the sense explained in Sect. 2.2.2), and indeed *if* they had complete information they would (hopefully) be in a position to fully explain the phenomenon. However, researchers seldom have full knowledge. In practice, what is needed is *enough* information to perform a *partial* decomposition (in the sense explained in Sect. 2.2.2) that is nonetheless interpretable and meaningful.

In other words, the recursive decomposition *may* be interpreted as a mechanism. But it is not the mechanism that does the explaining. It is the whole modelling procedure including the reasons for choosing some variables rather than others. Differently put, explaining a social phenomenon also involves specifying the reasons, or functions, of the causes, as we will discuss in the next section. It is of course an open question of how far a social mechanism can be recursively decomposed. In a systemic approach, where directed causal relations are replaced with multi-directed influence relations, recursivity would not be satisfied. This may indeed be a more faithful description of reality, but not necessarily a useful modelling strategy for explanation or intervention (see Russo 2010).

Third, not all models are explanatory, i.e. not all models in social science have to be structural. A good example are population projection models. Simplified models based on associations only can lead to better results than other more sophisticated ones. This of course opens Pandora's box concerning the relations between explanation, prediction and background knowledge. But we keep this box firmly closed in this paper.

4 Functions and Mechanisms

In the previous sections, we recalled the core idea behind structural-modelling explanations: causal models explain insofar as (1) they are able to perform a recursive decomposition and (2) this decomposition can be interpreted as a (social) mechanism. We stressed the importance of background knowledge to choose the variables and relations to include in the model. This amounts, to anticipate the arguments that will be developed in detail later, to specify the *function*, or role, or reason, played by such variables.

Function is however a philosophically loaded notion, on which debate is still vivid. It should be noted that the mechanism/functions view developed in this paper departs from the much broader functionalist perspective *à la* Radcliffe-Brown, Malinowski, or Merton. This originates in the works of Durkheim and Mauss, and has been further developed by Parsons and Merton (for a discussion, see e.g. Grawitz 2001). In this broader functionalist perspective, every social practice in a society must have (or must have had) a function in the society as a whole. Functionalism is concerned with the organizational structure of social systems: how the components are interrelated and what their function is in society. The analysis of functions is carried out in very general terms, independently of specific contexts.

We instead examine the notion of function in structural modelling, integrating it with the debates happening in the literature on mechanisms. Our arguments restrict the concept of function to specific mechanisms, and to its components, placed in a given context. The mechanisms used to explain a given phenomenon are represented by recursive decompositions, as explained in previous sections. Our notion of function and mechanism is, in a sense, closer to the approach adopted in analytical sociology, as they also emphasise the explanatory role of mechanisms for societal phenomena. But functionalism, in the much broader sense just alluded, is outside the scope of this paper.

In the following, we isolate two discussions of functions that are of interest to our purposes. The first is Illari and Williamson's discussion of function and organisation in mechanisms. The second is Franck's account of functions, which relies on the notion of 'functional architecture'. Both discussions help us clarify our position, which is presented in Sect. 4.3.

4.1 Isolated Descriptions and Role-Functions

The recent philosophical literature on mechanisms agrees that the individuation of the function and of the functioning of the mechanism is important. Functional individuation has however a ‘double’ facet. One is the *role* that either the whole mechanism or its components have in a given context, the other is the *description* of (the functioning of) a mechanism independently of any context. Illari and Williamson (2010, 283) say:

> The relevant sense in which a mechanism as a whole has a function depends on whether the mechanism is situated in a context, or considered alone. If it is situated in a context, then it can have a role-function. The mechanism for the pumping of the heart, for example, might itself have a role-function. This is so if that mechanism is to be understood within the context of explaining the circulation of the blood. That itself might further be understood in the context of keeping the animal alive, and so on. But if there is no reference to a containing system, a mechanism has only an isolated description, or characteristic activity. This is where functional descriptions top out.

Illari and Williamson (2010) then argue that these two aspects—the function of the mechanism and of its components on the one hand, and the ‘mere’ description of a mechanism on the other hand—are best cashed out by views developed respectively by Cummins and by Craver: functions are to be understood as ‘role-functions’ (Cummins 1975) and as ‘isolated descriptions’ (Craver 2001). Let us examine these in turn.

Isolated descriptions. For Craver, the function of a mechanism or of a component of it can be picked out in a contextual way or in a more isolated way (Craver 2001, 65):

> But this leaves it ambiguous whether the function is the capacity, described in isolation and simply ‘picked out’ by its contextual role, or, instead, the contextual role by virtue of which the capacity is picked out. A complete description of an item’s role would describe each of these … There is a difference, after all, between knowing that spark plugs produce sparks and knowing how that sparking is situated within the complex mechanisms of an engine.

In the social sciences, however, the usual practice is to describe mechanisms in a *contextual* way. For instance, alcohol consumption is considered in the mechanisms of socialisation; the researcher may then be interested in locating socialisation in the broader context of an individual’s social well-being, and so on. Therefore, it is of utmost importance to specify *what* mechanism is the object of study and in what *context* it is situated. In fact, a same factor can act differently, depending on the context considered: alcohol intake can also have a detrimental impact on physical and mental health, and is thus an important determinant of morbidity and mortality. We do not deny that isolated descriptions provide important insights about mechanisms, we just report that the practice in social science looks for descriptions of mechanisms situated in contexts, which leads us to the concept of ‘role-function’.

The function of the mechanism and of its components. Cummins (1975) develops the concept of ‘role-function’: the function of something is the role it plays in the overall behaviour (Cummins 1975, 762):

> x functions as a ϕ (or, the function of x in s is to ϕ) relative to an analytical account A of s’s capacity to ψ just in case x is capable of ϕing in s and A appropriately and adequately accounts for s’s capacity to ψ by, in part, appealing to the capacity of x to ϕ in s.

We could freely paraphrase Cummins in the case of the Baltic study as follows: in the complex mechanism explaining self-rated outcomes in the Baltic countries (for the specified periods of time), the function of alcohol consumption is to relieve stress. Data analysis, background knowledge, sociological and demographic theory support this interpretation of the function of alcohol consumption.

We should be careful, however, in distinguishing two issues. One question concerns the function that the *whole* mechanism has in a given context. The issue is to understand which mechanism(s) can account for the phenomenon we are interested in. This aspect, that mechanisms are always 'mechanisms *for* a phenomenon', has been emphasised by Stuart Glennan as early as 1996. Illari and Williamson illustrate with the example of the mechanism of the heart pumping blood in the context of blood circulation in the human body.

This way of identifying the role-function of a (whole) mechanism also happens in the kind of explanatory contexts we have been considering in this paper. In fact, structural-modelling explanation begins with the identification of the research question. For instance, suppose we observe a fall in fertility in Europe. Research would accordingly turn to the determinants of the fertility decrease. A structural-modelling explanation would then build a model in order to account for such a phenomenon. Such a model would be explanatory insofar as it represents a (plausible) mechanism of the fall of fertility. This would be the (role-)function of the mechanism *as a whole*.

Functional individuation does not stop here, however. Another question concerns the function of the components of a mechanism. The components of a mechanism have themselves a *role* in the production of the phenomenon. The question is now about the *functioning* of a mechanism, and we must understand what are the relevant components and what they do, namely what their function is. Here, the relevant concept of function is, again, Cummin's role-function. In fact, the task is to understand the role played by the components of the mechanism in order to produce the behaviour under examination. Illari and Williamson (2010, 285) explain this idea thus:

> Activities [*i.e.* what parts of mechanism do] are interesting because they often are the functions of the entities—an activity is what an entity does, or what two or more entities do together. Activities are individuated in a similar way to entities in the hierarchy of mechanisms. Activities are identified in terms of their contribution to the behaviour of the phenomenon to be explained.

Consider again the case of the heart pumping blood. We would now identify the key components in this mechanism and their activities, for instance the functioning of the ventricles, of the incoming and outgoing vessels, etc.

This type of functional individuation happens in social context too. Consider again the study on self-rated health in the Baltic countries in the early nineties (Gaumé and Wunsch 2010). Alcohol consumption had an impact on self-rated health, but its function or role (i.e. *why* Balts drink) seemed to lie in its stress-reducing and possibly in its socialisation effects. In other contexts, however, this stress-reducing role could be implemented by other empirical behaviours such as ritual dances, meditation, taking drugs, and so on, and the socialization role by attending church, participating in social and political activities, etc. According to the context, the same function can be implemented by different activities.

Structural-modelling explanation considered in this paper (implicitly) adopts the concept of role-function, applied to both the whole mechanism and to its components. Structural-modelling explanation operates in fact an analogous functional individuation. In structural modelling, the research question identifies the role-function of the whole mechanism, and the identification of the components of the mechanisms and of their role-function is done through the recursive decomposition, as explained earlier in Sect. 2. It is worth noting that the terminology used in the mechanisms literature (entity-activity, functional individuation, etc.) is usually not adopted in structural modelling literature. We hope to show in this paper that there is enough ground for borrowing these concepts from the mechanisms literature, because they can shed light on explanatory practices in social science too.

4.2 Functional Architectures

In numerous works, Robert Franck (1994, 1995, 2002, 2007) has stressed the role of the functional architecture of mechanisms. Franck (2002, 295) summarises his approach as follows:

> (1) Beginning with the systematic observation of certain properties of a given social system, (2) we infer the formal (conceptual) structure which is implied by those properties. (3) This formal structure, in turn, guides our study of the social mechanism which generates the observed properties. (4) The mechanism, once identified, either confirms the advanced formal structure, or indicates that we need to revise it.

Franck continues by stating that the role of empirical research is to describe and to analyse the properties of the social system under study. From these properties, a conceptual or theoretical structure is inferred, on the basis of the principles or functions putatively governing the properties. This inferred formal structure guides the investigation of the empirical mechanism generating the properties, i.e. it suggests possible empirical factors contributing to the outcome. Finally, the empirical mechanism confirms or disconfirms the value of the theory proposed.

Franck thus distinguishes between the theoretical functional structure proposed after examining the properties of the social process under study, and the empirical mechanism actually generating them. In other words, the theoretical model or functional architecture is not tied to a particular empirical mechanism—even though it is derived from empirical observations—and can be transferred to other social processes with similar properties, the theoretical properties or functions being possibly implemented by different functionally equivalent empirical means or causes and by their combination. It is however the functional architecture that is necessary to produce the social process and not the empirical mechanism (such as a causal model), which is contingent upon the context. In this sense, the functional architecture is to some extent generalizable while the empirical mechanism is not.

4.3 Functions in Structural Explanations

The discussion of Illari and Williamson and the approach of Franck set the ground for us to say how functions ought to be understood in the context of structural-modelling explanation. The meaning of function as ‘role-function’ and its importance for model building are pretty easily established. More debatable is instead the move of giving more importance to functions than to causes and thus privileging the ‘functional architecture’ over the ‘causal architecture’. A large part of this section is devoted to clarifying similarities with and divergences from Franck’s account.

Functions in structural explanations are role-functions. We agree with Illari and Williamson that role-functions are important for structural explanation. The concept of role-function, applied both to the mechanism as a whole and to its components, cashes out the meaning of function in the context we are interested in.

Structural explanations performed by recursive decomposition involve a global mechanism and organised sub-mechanisms. Such sub-mechanisms have role-functions too. In this approach, a conditional distribution is interpreted as a mechanism provided that it is supported by background knowledge. If the recursive decomposition has been suitably condensed, a conditioning variable in a mechanism affects the characteristics of the conditional distribution, such as the conditional expectation or variance, the fractiles or the tail of the conditional distribution etc. This reflects the role or function of the conceptual variable in the mechanism.

Functions, to be more precise, are the reasons why we choose some variables and relations rather than others. Thus, functions *already* belong to the methodology of structural-modelling explanation. In other words, they are not *new* elements; they are instead the theoretical underpinning of causes. Identifying the functions in a structural model means to strengthen the theoretical plausibility of the chosen recursive decomposition.

Functions help build the conceptual model. A consequence of this take on functions is that they are never 'established' with certainty but instead stay highly hypothetical in character. This derives from the hypothetical character of structural models in the first place. Structural models provide a mechanistic explanation of a phenomenon *to the best of our knowledge*. But this mechanism is not established forever, nor are the functions of the mechanism and of the sub-mechanisms. As Franck also stresses, the same functions can be implemented by different causes, and the same causes can have different functions.

We agree with Franck that causes should be incorporated into a putative mechanism only if we can point out their possible function, or rather role-function, in this mechanism. We should therefore avoid integrating into the putative mechanism variables associated with the outcome but for which background knowledge cannot specify their specific functions in the mechanism. We also agree with the view that a same function can be implemented by different means. Functions or roles relate to the conceptual level. We should distinguish between a conceptual model and its operational equivalent. Functions are not empirical variables, though they are derived from observation and implemented by empirical factors. We would like to point out however that the operational or empirical model, though based on observables, is also a theoretical construct: the abstract mechanism it reflects helps the scientist interpret social data. Actually, all models are abstract representations of reality. As Gudmund Hernes (1998, 78) writes:

> Reality presents itself to us, but we have to *represent* it in order to *make sense* of it. Mechanisms are the virtual reality of social scientists.

Causes, functions and contexts. Should an explanatory model be context-dependent or not? Franck gives an example from reverse engineering of an artificial heart, which is borrowed from de Callataÿ (2002, 108):

> If an artificial heart is grafted to a patient, this heart is an engineering model of the natural heart. This heart uses different pumping principles. It does not use the energy molecules produced by the body. […] But it still performs the heart's essential function.

We agree that the functions performed by the artificial heart are to some extent the same as some of those performed by the natural one, though the actual mechanisms are quite different. In fact, if the former fails, one will call upon an engineer to fix it and not upon a cardiologist; vice-versa, if one has an infarct, contacting a cardiologist might be preferable to asking help from an engineer. In other words, we think—diverging from Franck—that much research in social science is interested in *specific* contexts, in a particular place and time, and not in a relatively context-free general model. This is why 'role-functions' fit structural-modelling explanations better than 'isolated descriptions'.

Another reason why functions are most of the time context-sensitive is that social research is not conducted solely for knowledge but also for action. For example, we might want to understand or predict, and possibly act upon, the present economic situation of Greece, mortality in Russia, or the political evolution in Cuba. To give another example, one may wish to curb heavy drinking in the Baltic States due to its deleterious activity in the development of cardiovascular diseases, in addition to other causes such as tabacism and obesity. In this case, we are of course not interested in the functional equivalents of drinking as stress-relievers or socialisers, such as those given above, but on alcohol per se, that is on *what, why*, and *how* people drink in the Baltic states.

In social science, many factors and their interrelations are highly situation-related; it is doubtful that a comprehensive action-oriented explanation can be given with the help of a relatively context-free functional structure—conditional on the possibility of constructing one—rather than by a contextually-dependent empirical causal model. What one would possibly gain in *generality* would be lost in *specificity*.

A context-dependent causal explanation, as proposed in this paper, does not however mean it is necessarily singular (for a discussion of singular causal links when events are not repeated, see Abell 2011). Pierre Demeulenaere (2011b) has convincingly shown that an empirical social mechanism is often generalizable to a large extent because human actions have regular features, ensuring some stability or invariance at a certain time and place. This is firstly due to the fact that many human behaviours are grounded in institutions, rules and norms. People in France or Belgium usually drive on the right-hand side of the road because of the rules of the traffic law and because most people tend to follow these rules. Rules can vary of course in time or location, leading to other regularities. In Britain or in Japan one drives on the left, and Sweden has switched in 1967 from left-hand driving to right-hand driving, but this does not contradict the causal regularity feature.

According to Demeulenaere, there is a second reason why social mechanisms are relatively stable: human action is overwhelmingly predictable and agents engage in actions based on typical motives, anticipating typical actions of others that are expected under typical circumstances. Many actions are shaped by routine, even in the absence of institutions and rules, in anticipation of the typical expected reaction of others.

This does not mean of course that all individuals act in the same way. Actually, in many cases, the variance of behaviours is just as meaningful as their average. Rational behaviours are shaped by the institutional context and cultural norms, interrelations with other individuals, available information, resources, values, preferences, beliefs, motives, emotions (see for instance Boudon's theory of *ordinary rationality*, 1998 and 2011), and these factors can vary among individuals. For example, not all Balts are heavy drinkers; some are even teetotallers! Their drinking behaviour is related among other things to their age, gender, health status, ethnicity, education, and economic status (Gaumé and Wunsch 2010). Nevertheless, social regularities—i.e. common features of human action—are observed in many circumstances and back up the stability of the causal mechanism. Note that in our view, attributes such as gender or ethnicity can be part of the causal mechanism, contrary to the counterfactual /manipulation approach (see Russo et al. 2011, for a discussion).

One could therefore distinguish between the *functional architecture* of a social mechanism, as discussed above, which could possibly be generalised to some extent across contexts, and the *causal architecture* of a system, which would benefit from an intra-context relative stability. Depending upon the purpose of the research, requiring generality or specificity, one can focus the view either on the functional architecture or on the causal one.

5 Discussion and Conclusion

Structural modelling is a widespread practice in causal analysis, although the meaning of "structural" may vary considerably among scholars. It has been developed, defended, and discussed by a number of scholars in the social sciences and in philosophy. One peculiarity of the approach defended here is to provide an overarching view of modelling, that bridges the philosophical and statistical perspectives. Namely, structural models incorporate the whole set of assumptions, of testing procedures, and of background knowledge that is needed at each stage of the model building and model testing process; consequently, structural models are not just defined by a (set of) equations. Differently put, structurality lies in the whole model building process, which is not confined to a statistical translation of a pre-conceived theory. Background knowledge and invariance considerations play a central role and complement each other in the construction of the model. Conversely, the recursive decomposition and its possible interpretation in terms of mechanism is central to the explanatory aspect of a structural model. In earlier works we thoroughly discussed issues concerning the model building process and its explanatory features. In this paper, we focus on the role of the concept of "function" in such context.

In a structural modelling framework, the classification of variables into mediators, moderators, or confounding variables refers to the "role-function" of a variable in the functioning of a mechanism or sub-mechanism. In particular, the role of mediator, or of confounder, depends not only on the recursive decomposition (i.e. on the identified sub-mechanims), but also on the possible presence of "simplifying" assumptions (i.e. on the functioning of these sub-mechanisms). A moderator, or interaction effect, means that the effect of a causing variable, say Z, on an outcome variable, say Y, may depend on the values of other causing variables, say X. This is a property of the conditional distribution $p_{Y|X,Z}$ independently of the joint marginal distribution $p_{X,Z}$, and is not representable by a DAG. Moderation should accordingly be viewed in the framework of classifying different modes of interactions.

We examined the notion of function in structural modelling starting from the debates in the literature on mechanisms. We isolated two discussions that are of interest for our purposes. The first concerned the concept of function cashed out in terms of 'role-function' and of 'isolated description'; this is borrowed from the literature on mechanisms (in biology and neuroscience, mainly). The second concerned the concept of functional architecture; this is borrowed from Robert Franck's philosophy, that tackles social science more specifically.

Having examined these two alternatives, we offered an account of the concept of function in structural explanations. The overall conclusion is that the concept of role-function well suits structural modelling. It allows us to talk about the role, or function, a mechanism has in order to produce some behaviour or phenomenon we are interested in. It also allows us to talk about the role, or function, that the elements of a mechanism play. Concerning this second characterisation, this eventually means specifying the *reasons* to include some causes rather than others in the mechanism and specifying their role. In this sense, functions are the 'theoretical' underpinnings of the causes.

Understanding the relations between function, mechanism, and explanation is relevant because insisting on the importance of identifying the function of a mechanism and of its components (ought to) force the practicing social scientist to justify each step of the modelling procedure. There is no causal rabbit to extract from a statistical hat. There is instead a wise and critical use of background knowledge that helps the statistical machinery extract the best (possible) information from the available data.

Acknowledgments M. Mouchart gratefully acknowledges financial support from the "Inter-University Attraction Pole", phase VI (number P6/03) of the Belgian Government (Belgian Science Policy). F. Russo has been FWO Marie Curie Pegasus Fellow at the Vrije Universiteit Brussel during the academic year 2012–2013. We thank Robert Franck for his remarks on an earlier version of this paper and for commenting in greater detail his approach. Comments and suggestions from two anonymous referees helped us improve our arguments.

References

Abell, P. (2011). Singular mechanisms and Bayesian narratives. In P. Demeulenaere (Ed.), *Analytical sociology and social mechanisms*. Cambridge: Cambridge University Press.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in the History and Philosophy of the Biological and Biomedical Sciences, 36*, 421–441.

Biello, K. S., Murin, S., & Mattay, R. A. (2002). Epidemiology, etiology and prevention of lung cancer. *Clinics in Chest Medicine, 23*(1), 1–25.

Blalock, H. M. (1964). *Causal inferences in nonexperimental research*. Chapel Hill: University of North Carolina Press.

Boudon, R. (1967). *L' analyse mathématique des faits sociaux*. Paris: Librairie Plon.

Boudon, R. (1998). Social mechanisms without black boxes. In P. Hedström & R. Swedberg (Eds.), *Social mechanisms*. Cambridge: Cambridge University Press.

Boudon, R. (2011). Ordinary rationality: The core of analytical sociology. In P. Demeulenaere (Ed.), *Analytical sociology and social mechanisms*. Cambridge: Cambridge University Press.

Craver, C. (2001). Role functions, mechanisms and hierarchy. *Philosophy of Science, 68*(1), 53–74.

Craver, C. (2007). *Explaining the brain*. Oxford: University Press.

Cummins, R. (1975). Functional analysis. *Journal of Philosophy, 72*, 741–765.

de Callataÿ, A. (2002). Computer simulation methods to model macroeconomics. Chapter 5. In R. Franck (Ed.). *The explanatory power of models*. Dordrecht: Kluwer.

Demeulenaere, P. (Ed.). (2011). *Analytical sociology and social mechanisms*. Cambridge: Cambridge University Press.

Demeulenaere, P. (2011). Causal regularities, action and explanation. In P. Demeulenaere (Ed.), *Analytical sociology and social mechanisms*. Cambridge: Cambridge University Press.

Duncan, O. (1975). *Introduction to structural equation models*. New York: Academic Press.

Engle, R. F., Hendry, D. F., & Richard, J.-F. (1983). Exogeneity. *Econometrica, 51*(2), 277–304.

Fennell, D. J. (2011). The error term and its interpretation in structural models in econometrics. In P. Illari, F. Russo & J. Williamson (Eds.), *Causality in the sciences* (pp. 361–378). Oxford: Oxford University Press.

Florens, J.-P., & Mouchart, M. (1985). Conditioning in dynamic models. *Journal of Times Series Analysis, 53*(1), 15–35.

Florens, J.-P., Mouchart, M., & Rolin, J.-M. (1980). Réduction dans les expériences bayésienne séquentielles. *Cahiers du Centre d'Études de Recherche Opérationnelle, 22*(3–4), 353–362.

Franck, R. (Ed.). (1994). *Faut-il chercher aux causes une raison ?* Paris: Librairie Philosophique J. Vrin.

Franck, R. (Ed.). (1995). *Les sciences et la philosophie*. Paris: Librairie Philosophique J. Vrin.

Franck, R. (Ed.). (2002). *The explanatory power of models*. Dordrecht: Kluwer.

Franck, R. (2007). Peut-on accroître le pouvoir explicatif des modèles en économie?. In: A. Leroux & P. Livet (dir.) *Leçons de philosophie économique*, T.III (pp. 303–354). Economica, Paris.

Gambetta, D. (1998). Concatenations of mechanisms. In P. Hedström & R. Swedberg (Eds.), *Social mechanisms*. Cambridge: Cambridge University Press.

Gaumé, C., & Wunsch, G. (2010). Self-rated health in the Baltic countries, 1994–1999. *European Journal of Population, 26*(4), 435–457.

Glennan, S. (1996). Mechanisms and the nature of causation. *Erkenntnis, 44*, 49–71.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*(Supplement), 342–353.

Grawitz, M. (2001). *Méthodes des sciences sociales* (11ème édition). Paris: Dalloz.

Hedström, P., & Swedberg, R. (Eds.). (1998). *Social mechanisms*. New York: Cambridge University Press.

Heckman, J. (2008). Econometric causality. *International Statistical Review, 76*, 1–27.

Hernes, G. (1998). Real virtuality. In P. Hedström & R. Swedberg (Eds.), *Social mechanisms*. Cambridge: Cambridge University Press.

Hood, W. C., Koopmans T. C. (Eds.). (1953). *Studies in econometric methods, Cowles commission for research in economics, monograph no 14*. New-York: Wiley.

Hoover, K. (2011). Counterfactuals and causal structure. In P. Illari, F. Russo & J. Williamson (Eds.), *Causality in the sciences* (pp. 338–360). Oxford: Oxford University Press.

Illari P. M., Williamson J. (2010). Function and organization: Comparing the mechanisms of protein synthesis and natural selection. *Studies in History and Philosophy of Science, Part C, 41*(3), 279–291.

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

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

Koopmans, T. C. (Ed.). (1950a). *Statistical inference in dynamic economic models, Cowles foundation monograph 10*. New-York: Wiley.

Koopmans T. C. (1950b). When is an equation system complete for statistical purposes? In: Koopmans T. C. (Ed.), *Statistical inference in dynamic economic models* (pp. 393–490). New-York: Wiley.

Koopmans T. C. (1953). *Identification problems in economic model construction, chap. II* (pp. 27–48). In W. C. Hood & T. C. Koopmans (Eds.), *Studies in econometric methods, Cowles commission for research in economics, monograph no 14*. New-York: Wiley.

Laaksonen, M., Rahkonen, O., Karvonen, S., & Lahelma, E. (2005). Socioeconomic status and smoking. *European Journal of Public Health, 15*(3), 262–269.

Little, D. (1991). *Varieties of social explanations: An introduction to the philosophy of social science*. Boulder: Westview Press.

Little, D. (2011). Causal mechanisms in the social realm, chap. 13. In P. Illari, F. Russo & J. Williamson (Eds.), *Causality in the sciences*. Oxford: Oxford University Press.

Machamer, P., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.

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

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

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

Mouchart, M., Russo, F., & Wunsch, G. (2010). Inferring causal relations by modelling structures. *Statistica, LXX*(4), 411–432.

Mouchart, M., & Vandresse, M. (2007). Bargaining power and market segmentation in freight transport. *Journal of Applied Econometrics, 22*, 1295–1313.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Pearl, J. (2011). The structural theory of causation. In P. Illari, F. Russo & J. Williamson (Eds.), *Causality in the Sciences* (pp. 697–727). Oxford: Oxford University Press.

Russo, F. (2009). *Causality and causal modelling in the social sciences. Measuring variations*. New York: Springer.

Russo, F. (2010). Are causal analysis and system analysis compatible approaches? *International Studies in Philosophy of Science, 24*(1), 67–90.

Russo, F. (2011a). Correlational data, causal hypotheses, and validity. *Journal for General Philosophy of Science, 42*(1), 85–107.

Russo, F. (2011b). Explaining causal modelling. Or, what a causal model ought to explain. In M. D'Agostino, G. Giorello, F. Laudisa, T. Pievani & C. Sinigaglia (Eds.), *New essays in logic and philosophy of science* (SILF Series, Vol. I, pp. 347–361). London: College Publications.

Russo, F., Wunsch, G., & Mouchart, M. (2011). Inferring causality through counterfactuals in observational studies—Some epistemological issues. *Bulletin of Sociological Methodology/Bulletin de Méthodologie sociologique, 111*, 43–64.

Saracci, R. (2006). Asbestos and lung cancer: An analysis of the epidemiological evidence on the asbestos-smoking interaction. *International Journal of Cancer, 20*(3), 323–331.

The Social Issue Research Centre. (1998). Social and cultural aspects of drinking. A report to the Amsterdam Group. http://www.drugslibrary.stir.ac.uk/documents/social_drinking, accessed 26 February 2013.

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

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

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.

Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics, 5*(3), 161–215.

Ylikoski, P. (2011). Social mechanisms and explanatory relevance. In P. Demeulenaere (Ed.), *Analytical sociology and social mechanisms.* Cambridge: Cambridge University Press.