Explaining causal modelling. Or, what a causal model ought to explain

Federica Russo
Philosophy, Louvain & Kent

Abstract One of the goals of the social sciences is to explain social phenomena. In this explanatory enterprise, causal relations, established by means of the so-called 'causal models', play a major role. This paper advances the view that causal models, by modelling causal mechanisms, (ought to) provide an explanation of social phenomena and should be seen as a model of explanation.[1]

1. Introduction

One of the goals of the social sciences is to understand social phenomena, that is to exhibit the mechanism underlying and bringing them about. This task goes beyond description: to exhibit this mechanism requires identifying *causal* relations between variables of interest. In quantitative social research, causal models are used to provide such explanations of social phenomena. This paper investigates whether causal models can be seen as *models of explanation*, and argues that causal modelling, by modelling causal mechanisms, provides (or ought to provide) genuine causal explanations and should be considered as a model of explanation, notably a *hypothetico-deductive* model of explanation.

The paper is organised as follows. *Section two* presents and explains what a causal model is and what it is supposed to do. *Section three* analyses how the terms 'explanation', 'explanatory', or 'explain' are used in causal models. *Section four* advances the view that causal modelling ought to be the modelling of causal mechanisms. *Section five* builds on the results of the previous sections and argues that causal models are a *model of explanation*, in particular, they are *hypothetico-deductive models*, where the H-D structure of the explanation is given by the H-D methodology of causal models. Finally, *section six* compares causal modelling with other models of explanation—notably, the deductive-nomological, statistical-relevance, the causal-mechanical, and the manipulationist model—and shows why those models of explanation are not fully satisfactory in social science and particularly in quantitative social science. This enables us to highlight what causal modelling offers over and above traditional models of explanation.

2. Causal modelling

A causal model consists of a set of mathematical equations (also called structural equations) and/or of a graph laying down the hypothesised causal structure pictorially.[2] More technical and precise definitions of causal models are of course possible. For one account, and for detailed examples accessible to a non-specialised audience, see Russo

[1] A shorter version of this paper is going to appear in the Proceedings of the SILFS Conference, Milan 8-10 October 2007.

[2] A number of causal models rely or employ structural equations in an essential manner—for instance, covariance structure models or multilevel models—but others do not—for instance, counterfactual models. However, although counterfactual models are rightly called 'causal' because they seek to measure the average causal effect of a treatment or intervention, surely they substantially differ from structural equation models in that they do not aim at modelling the causal mechanism, which, as we shall see later, is an essential feature of structural equation-type models.

(2008, ch.3), Russo et al (2006) and Mouchart et al (2007), and references therein.[3] For the purpose of the present paper we can keep technicalities to a minimum. Typically, both equations and graphs are employed. However, one can start by writing the equations and subsequently drawing the graph or the other way round.

Causal models normally consist of sets of equations, indicating a web of causal and non-causal relations among the variables of interest. To illustrate, consider the study on health care system and mortality (Lopez-Ríos et al. 1992). The causal model consists of the graph in figure 1 and of the following equations.

$$Y = \beta_2 X_2 + \beta_4 X_4 + \varepsilon_1$$
$$X_2 = \beta_{12} X_1 + \varepsilon_2$$
$$X_3 = \beta_{13} X_1 + \varepsilon_3$$
$$X_4 = \beta_{34} X_3 + \beta_{54} X_5 + \varepsilon_4$$

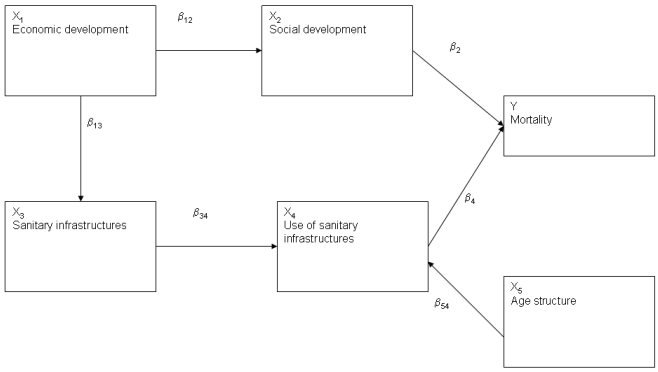

FIG. 1 Health system and mortality

In this model, each variable is regressed on its immediate ancestors. X_1 and X_5 are exogenous, thus meaning that they are the causes of mortality and the other variables are intervening variables, having a causal role just as effects of X_1 and X_5 but not on their own. In plain English, the equations state that regional mortality is causally determined by social and economic development; sanitary infrastructures affect mortality through the use of infrastructures, but the use of infrastructures depends on economic development, so it is not an exogenous cause; use of sanitary infrastructures, in turn, is an effect of the age of individuals and of the variable 'sanitary infrastructure', so again it is not an exogenous cause, but it has causal impact on mortality only through other factors.

An important feature of causal models is that they rest on a number of assumptions, some of which are merely statistical and others have instead causal import. Among the statistical assumptions we find, for instance, linearity and normality, non-measurement error and non-correlation of error terms. Those are standard statistical assumptions also made in associational models. However, causal models are provided with a much richer apparatus that allows their causal interpretation. In this apparatus we find background knowledge of the causal context, the conceptual hypothesis, a number of extra-statistical assumptions and of causal assumptions. Among extra-statistical assumptions we can list

[3] Disagreement arises in causal modelling as to whether structural equations and directed acyclic graphs convey exactly the same information. Partisans of the former approach tend to give a negative answer, because in graphical models some assumptions are relaxed and relations between variables are not expressed with the mathematical precision of structural equations. On the other hand, supporters of graphical models, such as Bayesian nets, maintain that this formalism indeed provides a simplification with respect to structural equation modelling without loss of any relevant information.

the direction of time, causal asymmetry, causal priority, causal ordering, and the deterministic structure of the causal relation. Causal assumptions include: structure of the causal relation (separability), covariate sufficiency, no-confounding, non-causality of error terms, stability, and invariance. A large part of causal models used in social science, unlike associational models, use a hypothetico-deductive methodology, according to which causal hypotheses are confirmed or disconfirmed depending on the results of tests and on whether they are congruent with background knowledge.[4] For a detailed account of the features of causal models and for a comparison with associational models see Russo (2008, ch.3); some of these features will nonetheless be discussed later.

An important characteristic of causal models is that causal relations are *statistically modelled*. This aspect deserves attention because influential philosophers of causality such as Wesley Salmon believed that aleatory causality will give a better understanding than statistical causality even in the social domain (Salmon 1990b). On the one hand, aleatory causality bestows emphasis upon the *physical* mechanisms of causality, primarily uses concepts such us 'process' and 'interaction', and appeals to laws of nature such as the conservation of energy or momentum. In the Salmon-Dowe (Salmon 1998, Dowe 2000) theory, causal processes are the key because they provide the link between the causes and the effects; causal processes intersect with one another in interactive forks, and in this interaction they are both modified and changes persist in those processes after the point of intersection. Causal processes and interactions are physical structures and their properties cannot be characterised in terms of probability values alone.

On the other hand, statistical causality puts emphasis upon constant conjunction and statistical regularity and uses, above all, concepts such as statistical relevance, comparison of conditional probabilities, or screening-off relations. Those concepts can be defined solely in terms of statistical terms, without resorting to any physical notion. According to the received view, statistical regularities are the 'symptoms' of causal relations. The conjunctive fork (Reichenbach 1956) gives the probabilistic structure of the causal relation and the screening-off relation alerts us about possible situations in which, given the correlation between two events *A* and *B*, a third event *C* may be responsible for their correlation.

But what do causal models do? Causal models *model* the properties of a social system. In particular, they model the relations between the properties or characteristics of the system, which are represented by variables.[5] In causal modelling, to model the properties of a social system means to give the scheme, or the skeleton, of how these properties relate to each other. In other words, the causal model *models* the causal mechanism governing the social system. However, this causal mechanism is not modelled in terms of spatio-temporal processes and interactions à la Salmon but is statistically modelled. This means that the concepts typical of statistical causality do help in identifying the types of relationships that hold among the variables of interest. In particular, causal models seek to uncover stable *variational* relations between the characteristics of the system. It is worth-noting that the received view of statistical causality, a heritage of Hume and represented, for instance, in the works of Suppes (1970), Eells (1991), and Cartwright (1989), emphasises the role of statistical *regularities* for assessing causality.

---

[4] It is worth noting that causal models can also be used in an inductive way, e.g. data mining. This is, for instance, the approach of Spirtes et al. (1993). Inductivist approaches claim that causal relations can be bootstrapped from data without the burden of extra-statistical and causal assumptions made in their hypothetico-deductive counterparts. Unfortunately, it goes far beyond the goal of this paper to discuss the success of inductive causal models. Consequently the scope will be limited to causal models that employ a hypothetico-deductive methodology.

[5] By 'social system' I simply mean a given population and, in the case study mentioned before, this is the Spanish population that have access to the health infrastructures. 'Population' has to be understood here in the statistical sense, that is as a set of units, those units being individuals, households, firms, etc.

In Russo (2006 and 2008) I challenge this view and argue that, instead, probabilistic theories of causality as well as causal modelling are governed by a rationale of variation, not of regularity. In a nutshell, the rationale of variation states that causal models measure and test joint *variations* between variables of interest, not regular sequences of events. To be sure, causal hypotheses are variational claims—that is they hypothesise how the effect would vary according to variations in the cause—and empirical testing aims at establishing whether variations are causal (rather than chancy), not whether regularities are causal. Of course, to ensure that variation can be interpreted causally, we have to impose further *constraints*, and a condition of invariance is required in order to interpret variations causally. It is worth noting that such invariance condition is not a condition of regular occurrence of events, but of stability of the model's parameters across different environments. In other words, variations among variables of interest will be deemed causal if the parameters are sufficiently stable across different environments. In fact, the invariance condition ensures that accidental and spurious relations be ruled out.

3. Explanation in structural equations

Consider now a simple form of a structural equation:

$$Y = \beta X + \varepsilon$$

where Y represents the putative effect, X represents the putative cause, β is a parameter quantifying the causal effect of X on Y, and ε represents errors or unmeasured factors. The scientific literature is not very homogeneous as to the vocabulary used. Y and X are called in a variety of ways depending on the specific discipline. Statistical textbooks will normally refer to X and Y as the independent and dependent variables, or as the explanatory and response variables, respectively; the econometric literature usually talks about exogenous and endogenous variables; the epidemiological literature spell them out in terms of exposure and disease, etc. Let us focus on the explanatory-response vocabulary, which perhaps constitutes the background of all disciplines that use causal models. In this case terminology is quite explicit: the *X*s supposedly *explain Y*. But what do exactly the *X*s explain? And how?

With much disappointment to the philosophers, in the scientific literature there is no explication of the terms 'explanatory', 'explanation', 'explain'. Intuitively, the *X*s explain *Y* in the sense that they 'account for' *Y*, namely the *X*s are relevant causal factors that operate in the causal mechanism, which is formalized by the equations and the graph. Needless to say, this is a very unsophisticated explication of 'explanation', yet intuitively clear. Let us leave aside, for the time being, the issue of what a good explanation is and of what causal modelling would offer over and above alternative models of explanation, and let us focus on what explanation in causal modelling consists of.

In causal modelling, the goal is to explain the *variability* in *Y*. Structural equations can be interpreted thus: variations in *X*s explain variations in *Y*, or variations in *X*s produce a variability in *Y*. Therefore, as long as we can control variations in *X*s we can also predict how *Y* will accordingly vary. The βs quantify the causal impact, or the direct causal effect of each of the *X*s on *Y*. So one can suggest that the more variability we can account for, the higher the explanatory power of the causal model. But how is this explanation 'quantified'? The (statistical) answer lies in the coefficient of determination r^2, which is the square of the correlation coefficient r and is a statistic used to determine how well a regression fits the data. It represents the fraction of variability in *Y* that can be explained by the variability in *X*s; thus r^2 indicates how much of the total variation in *Y* can be accounted for by the regression function.

However, this statistical answer is insufficient. This is for three reasons. The *first* is that r^2 just measures the goodness of fit, not the validity of the model, and a fortiori it does not say how well the model *explains* a given phenomenon. So r^2 gives us an idea of whether the variability in the effect is accounted for, and to what extent, by the covariates we chose to include in the model. But, and here is the *second* reason, the coefficient of determination does not give any *theoretical* motive for that. *Third*, the coefficient of determination will give us an accurate quantification of the amount of variance of *Y* explained by the *X*s only if the assumptions are correct. For instance, r^2 can be small either because *X* has only a small impact on *Y* (controlling for appropriate covariates) and/or because the relation between *X* and *Y* is not linear.

Instead, a more satisfactory (philosophical) answer lies in the specific features of causal models. Let us now revert to the assumptions of causal model and examine their explanatory import. Among the features of causal models listed in the previous section, those having explanatory import are two causal assumptions—notably, covariate sufficiency and no-confounding—and background knowledge.

Covariate sufficiency assumes that the independent variables are direct causes of the dependent variable, and that these are all the variables needed to account for the variation of the dependent variable. No-confounding then plays a complementary role in assuming that all other factors liable to screen-off the causal variables are ruled out. Those two together convey the idea that the causal model includes all and only the factors that are necessary to explain the variability of *Y*. Those assumptions rely on the hypothesis of the closure of the system, namely causal modelling assumes, so to speak, that we can isolate a mechanism within the larger social system under consideration, and that this mechanism is not subject to external influences. Thus, we can account for *Y*—that is for its variability—just relying on the factors we decided to include. This is indeed a strong assumption but that is the only way to go in order to avoid an *ad infinitum* regression hunting for more and more ancestral causes, and in order to exclude that everything influences everything else in the system, thus making impossible to identify the causal relations to intervene upon. Covariate sufficiency and no-confounding also highly depend on which variables we choose to include in the causal model. This choice, in turn, depends on background knowledge. But what is background knowledge in the first place?

The notion of background knowledge belongs to most quoted and least explicated concepts in causal modelling. Anything could fit in it. Unfortunately, if anything can be background knowledge, we lack a sensible criterion to say when and why the covariates contribute toward the explanation of the response variable. Background knowledge may include: (i) similar evidence about the same putative mechanism, (ii) general knowledge about the socio-political context, (iii) knowledge of the physical-biological-physiological mechanism, (iv) use of similar or different methodologies or of data. Different studies normally consider different populations. Differences can accordingly concern time, geographic location, basic demographic characteristics, etc. Background knowledge has to be used to justify the choice of the explanatory variables. This justification relies on the different aspects mentioned above.

Let us go through the case study presented in section 2 to show the explanatory role of covariate sufficiency, no-confounding, and background knowledge.

An example: Health care system and mortality

In 1992 the *European Journal of Population* published the results of a causal analysis on the incidence of the health system on regional mortality in the adult Spanish population (Lopez-Ríos et al. 1992). In this study, knowledge of the causal context is given by knowledge of the political, economic and social situation in Spain over the decades 1970-

1980, which led to a low mortality rate at the time of the study. Knowledge of the socio-political context sheds light on the modelling strategy of this study. In fact, previous studies in demography and medical geography examined the incidence of the health system on regional mortality, coming to the conclusion that regional differences in mortality could not possibly be explained by regional differences in the health system. However, Spain met deep socio-economic changes in the mid-Seventies, and consequently public policies in that period simultaneously tried to intervene on improving the social and the economic situation.

The causal model developed in this study highly depends on this background knowledge: mortality is influenced by the health system which is in turn influenced by social and economic development. It is this background that explains the choice of distinguishing the supply and demand of medical care, unlike the majority of similar ecological studies. Concerning the extra-statistical assumptions, causal priority and causal ordering are needed. In fact, a temporal gap is assumed between the putative causes (health system, economic and social development) and the putative effect (mortality). It is worth-noting that this is a model at the contextual level (administrative units) and not at the individual level: the results cannot be applied to individuals but to geographical areas. So, in order to avoid fallacious ecological correlations, the causes of regional mortality have been chosen at the same level, that is at the aggregate and not individual level.

Another important aspect of this study is the control of the variables 'social development' and 'age structure' that could confound the relation between the sanitary infrastructure and mortality by influencing the use of sanitary infrastructures. Here background knowledge plays a fundamental explanatory role. In fact, the goal of the study is to explain differences in regional mortality, and the choice of including economic and social development comes from knowledge of socio-political context Spain met in that time-lag. Furthermore, the choice of considering the distinction between the supply and demand side of the sanitary infrastructure offers other insights to explain differences in regional mortality. It is worth noting, however, that in analysing the relation between general mortality and main causes of death, the result is, unexpectedly, that the higher the use of sanitary infrastructure, the higher the mortality level. This result, according to the authors, can be explained in two ways. First, the indicators used in the study reflect curative and not preventive treatments. Second, an important covariate is possibly missing—that is a factor that simultaneously influences the use of sanitary infrastructure and mortality. Differently put, the assumption of covariate sufficiency does probably not hold in this case and the explanation is therefore incomplete.

4. Modelling causal mechanisms

So far, I argued that causal models attempt to explain the *variability* of the effect variable by means of appropriate covariates. I also argued that the explanatory import is given by specific causal assumptions made in causal modelling—notably, covariate sufficiency and no-confounding—and by background knowledge. This philosophical answer complements the statistical answer according to which the coefficient of determination quantifies the explanatory power of a causal model. Let us now go back to what causal models do. Earlier, I briefly put forward the idea that causal models model the properties of a system and that we could conceive of them as the scheme or skeleton of the causal mechanism governing the causal system under investigation. It is now time to develop this idea further.

The notion of mechanism is often evoked both in the scientific and philosophical literature. No account seems to attract an unanimous consensus, yet various

characterisations stick to the physical notions of process and interaction. That is to say, the most widespread conception sees mechanisms as made of physical processes, interactions, and of physical elements, somehow assembled together to behave like a gear. For instance, Little (2004) opposes Humean causality, that sees causation as mere regularity, to a realist view, that sees causal mechanisms and causal powers as fundamental. According to the realist, says Little, "a mechanism is a sequence of events or conditions governed by lawlike regularities leading from the explanans to the explanandum". Partisans of this view are obviously Salmon (1984), for he believes that causal processes, interactions and laws give the causal *mechanisms* by which the world works and that to understand why something happens we have to disclose the mechanism that brought it about, and Dupré and Cartwright (1988) who, in the same vein, argue that discovering causal relations requires substantial knowledge of the capacities of things or events—i.e., their power to bring about effect. Likewise, Bunge (2004) ultimately reduces mechanisms to physical process that interconnect with one another as is the case in most biosystems or physical systems. Also, in his account mechanisms are governed by causal laws.

Unfortunately, this view doesn't fit the case of the social sciences for two reasons. *First*, if the causal model only involves socio-economic-demographic variables, we cannot identify causal mechanisms in terms of *physical* processes and interactions (at least at that level of description). *Second*, if the causal model involves both social and biological variables, the causal mechanism will not be able to account for the 'social' part. Let me explain these two reasons further. The problem is that, in social contexts, mechanisms are not always, or not necessarily, made of physical processes and interactions. For instance, Lopez-Rios et al. (1992) modelled the relations between regional mortality in Spain and the use of sanitary infrastructure. This model does not involve *physical* processes and interactions. The mechanism described by the authors rather explains the behaviour of a social system in terms of the relations between some of its properties. These properties, however, do no necessarily have 'physical' reality as they might just be conceptual constructs, as for instance economic and social development. Consequently, the process leading, say, from economic development to mortality through the use of sanitary infrastructures does not correspond to a *physical* process, such as billiard balls colliding, but rather is our conceptualisation and schematisation of a much more complex reality.

Another difficulty is that if causal mechanisms are governed by causal laws, it is unclear where these laws come from, and if they come from causal modelling itself, then this leads to a vicious circle. So if we want to keep a physical notion of mechanism the price to pay is quite high—we would have to renounce to causal mechanisms in the social domain. We are not forced to this solution, though, if we are prepared to accept a wider concept of causal mechanism, in particular one that is based on causal modelling.

In a nutshell, causal modelling is, and ought to be, the *modelling of mechanisms*. A statistical characterization of mechanisms, along with a rationale of variation, is what mediates our epistemic access to causal relations. The net gain of this perspective is a non-physical characterization of causal mechanisms. In fact, mechanisms would then have observable components (corresponding to observable variables) and the only non-observed parts of causal mechanisms would be nodes representing latent variables. However, far from giving causal mechanisms a mysterious or epistemically inaccessible appearance, latent variables ought to be introduced to facilitate the interpretation of complex causal structures.[6]

[6] A similar view that emphasises the central explanatory role of mechanisms is advanced by Franck (2007). However, Franck's approach differs from mine in that it goes further in claiming that the modelling of a social mechanism ought to be completed with the modelling of the functions of the same mechanism.

Most importantly, the modelling of mechanisms ought to rely on the rationale of variation rather than on the rationale of regularity: causal mechanisms are made of *variational* relations rather than regular relations. The components of the causal mechanisms are arranged depending on what variations hold. Agreed, those variational relations happen to be regular (or at least regular enough), but this depends on the fact that causal modelling analyses large (enough) data sets. Furthermore, regularity does not seem to be successful in constructing causal mechanisms, for the Humean view and the realist view eventually collapse in the same tenet. In fact, according to the realist, the sequence of events in causal mechanisms is, in the ultimate analysis, governed by a lawlike *regularity*.

Such characterisation of causal mechanism allows us to incorporate in the causal model both socio-demo-political variables and biological variables. It goes without saying that pathways in such a mixed mechanism have to be made explicit as there isn't homogeneity at the ontological level. Health variables do not cause changes in social variables (or vice-versa) *as such*. Socioeconomic status influences one's health through the possibility of accessing some sanitary infrastructures, but not directly. Arguably, the social sciences are interested in identifying causal mechanisms that involve different types of variables—this interdisciplinary stance is also a perspective undertaken in epidemiology (see for instance, Susser and Susser (1996)).

To sum up, if causal models do not model the mechanism underlying the phenomenon being investigated, they lack explanatory power. To see why it is so, let us compare them with associational models. Associational models only investigate statistical associations between variables, but no causal interpretation is allowed for the parameters. This is due to several reasons. *First*, associational models don't have the rich apparatus of statistical, extra-statistical, and causal assumptions as causal models do—associational models are normally equipped just with statistical assumptions. *Second*, they do not employ a hypothetico-deductive methodology—there is no formulation of the causal hypothesis because it is not their goal to confirm or disconfirm hypotheses. Thus modelling mechanisms, that is identifying the causal interrelations between the variables of interest, becomes a necessary condition for the explanatory power of causal models. The question then arises as to what kind of formal structure such explanation should have—this issue will be tackled in section 5.

An example: child survival in developing countries

To illustrate this notion of causal mechanism in social science I shall use a case study from demography, notably, the model for child survival in developing countries put forward by Mosley and Chen (1984). Infant mortality, especially in developing countries, is and has been the object of numerous studies. In the late Seventies, the eminent demographer Caldwell (1979) hypothesised that maternal education played a major role in explaining child mortality. He supported this claim by showing that other socio-economic factors traditionally thought to be the main causal factors—i.e., mother's place of residence, husband's occupation and education, and type of marriage (monogamy or polygamy)—did not explain, altogether, as much variation in the response variable as maternal education, alone, did. Caldwell justified this result by saying that educated African women were likely to better communicate with doctors because they were less fatalistic about illness, and, more generally, in African households education of women greatly changed the traditional balance of familiar relationships with significant effect on child care. Caldwell's model for child survival has been the accepted paradigm for many years. Caldwell's merit was to point to a factor that better explained the variation in the response variable—that is he developed a model with a better coefficient of determination—but he did not push his study as far as spelling out the *causal mechanism*

behind this. Mosley and Chen built upon Caldwell's results and put forward an analytical framework in which a more complex and mixed causal mechanism was used to explain the variation in child survival.

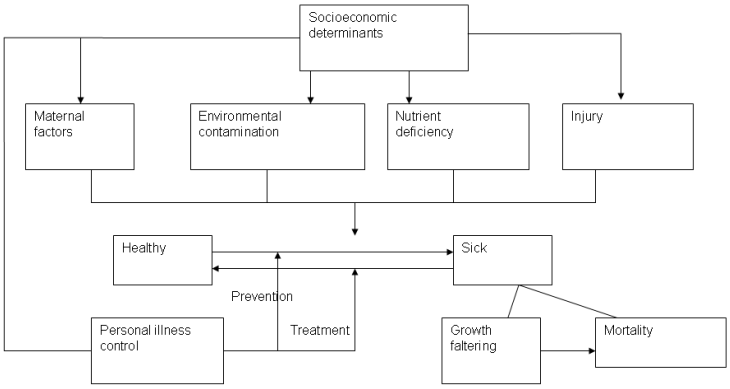

FIG. 2 Child survival in developing countries

In this graph a biological mechanism and a socio-economic mechanism are combined in order to explain *how and why* mother's education influences child survival. The answer to the how-question is given by the all the possible paths in the causal mechanism. The answer to the why-question is given by some non-mathematized knowledge, viz. by background knowledge that informs both the model-building stage and interpretation of results. Building a faithful causal mechanism is, needless to say, essential for our cognitive endeavour to understand society. Nonetheless, the action-oriented dimension cannot be underestimated. The clear and exact specification of causal mechanisms is important for policy-making purposes, since it indicates different levels at which interventions are possible or feasible. Good modelling of causal mechanisms ought to inform us on the 'means' we have to intervene upon.

5. Causal modelling as a model of explanation

The two previous sections argued that some specific features of causal models have explanatory import and that if causal models can be successful at all in the enterprise of explaining social phenomena, this is because they model causal mechanisms. I will now advance the view that causal models are models of explanation, in particular, they are *hypothetico-deductive* models of explanation.

The formal structure of the explanation is given by the hypothetico-deductive character of model-building and model-testing of causal modelling. Simply put, hypothetico-deductivism is the view according to which scientists first formulate hypotheses and then test them by seeing whether or not the consequences derived from the hypotheses obtain. Popper (1959), who first developed the H-D methodology, was motivated by the need of providing a scientific theory in a non-inductive way. However, in causal modelling, hypothetico-deductivism takes a slightly different facet specifically concerning deduction, but does borrow from the Popperian account the primary role of the hypothesis-formulation stage. I shall get back to this point shortly.

According to the H-D methodology, model building and model testing essentially involve three stages: 1. formulate the causal hypothesis; 2. build the statistical model; 3.

draw consequences to conclude to the empirical validity or invalidity of the causal hypothesis.

The hypothesis to put forward for empirical testing does not come from a tabula rasa, but emerges within a causal context, namely from background theories, from knowledge concerning the phenomenon at stake, and from preliminary analyses of data. This causal hypothesis, which is also called the 'conceptual hypothesis', is not analysable *a priori*, however: its validity is not testable by a logico-linguistic analysis of concepts involved therein. On the contrary, to test the validity of the causal hypothesis requires building a statistical model, and then drawing consequences from the hypothesis. For instance, in the case study on health systems and mortality presented earlier, background knowledge and previous studies indicate the hypothesis to formulate and the consequences that researchers draw concern exactly what they expect to be accepted or rejected at the testing stage. The estimation of the statistical model and hypothesis testing will allow us to conclude to the empirical validity or invalidity of the causal hypothesis.

If the model is correctly estimated and fits the data, the hypothesized causal link is accepted—provided that it is congruent with background knowledge. The hypothetico-deductive structure of causal modelling is thus apparent: a causal hypothesis is first formulated and *then* put forward for empirical testing. That it to say, the causal hypothesis is *not directly inferred* from the data gathered, as is the case with inductive strategies, but accepted or rejected depending on the results of tests.

As anticipated above, hypothetico-deductivism in causal modelling does not involve deductions *strictu sensu*, but involves a weaker inferential step of 'drawing consequences' from the hypothesis. That is to say, once the causal hypothesis is formulated out of the observation of meaningful co-variations between the putative cause and the putative effect and out of background knowledge, we do not require data to be *implied* by the hypothesis but just that data conform to it. Here, 'conform' means that the selected indicators *adequately* represent the conceptual variables[7] appearing in the causal hypothesis. Thus, this way of validating the causal hypothesis is not, strictly speaking, a matter of deduction, but surely is, broadly speaking, a deductive procedure. More precisely, it is a *hypothetico*-deductive procedure insofar as it goes the opposite direction of inductive methodologies: not from rough data to theory, but from theories to data, so to speak.[8]

To sum up, a causal model attempts to *explain* a given social phenomenon—in particular, the variability of the effect variable *Y*—by means of number of explanatory variables *X* and the explanatory procedure is given exactly by the hypothetico-deductive methodology of causal models. How do we evaluate the goodness or the success of the explanation then? We have seen before that the coefficient of determination is insufficient to provide such an answer, which instead lies in the peculiar features of causal models. Statistical tests, notably invariance and stability tests, provide the accuracy of measurements but alone cannot guarantee the explanatory goodness of the causal hypothesis. In fact, non-sense correlations, such as the monotonic increase of both bread prices in England and sea-level in Venice, may well turn out to be stable or invariant and yet not causal nor explanatory at all. The goodness of an explanation cannot be assessed on statistical grounds *alone*—the story also has to be coherent with background

[7] A conceptual variable is a variable that cannot be measured directly but from some 'indicators'. For instance, socio-economic status can be measured by taking into account income and years of schooling.

[8] For a discussion of the H-D method at work in the social sciences see Russo (2008, ch.3.2), and also Cartwright (2007, ch.2). Cartwright, as many others both in the philosophical and scientific literature, calls the methodology of causal models hypothetico-deductive but she also warns us about the weaker form of deductivism hereby involved.

knowledge and theories previously established, and has to be of practical utility for intervening on the phenomenon.

Thus, the problem of the goodness of explanation is mainly a problem of internal validity, with the caveat that, among various threats, coherence with the background plays a major role.[9] This, however, makes explanation highly context-relative simply because the causal model itself is highly context-relative. This could be seen as a virtue, as restricting the scope leads to more accurate explanations. But obviously this situation raises the problem of generalising results—that is the external validity of the causal model. It goes far beyond the scope of the present work to advance the criteria that allow the generalisation to a different population and/or different time.

Hypothetico-deductive explanations also exhibit a flexibility rarely found in other models. First, they allow a *va et vient* between established theories and establishing theories. Established scientific theories are (and ought to be) used to formulate the causal hypothesis and to evaluate the plausibility of results on theoretical grounds. But causal models also participate in establishing new theories by generalising results of single studies. This reflects the idea that science is far from being monolithic, discovering immutable and eternal truths. If the model fits the data, the relations are sufficiently invariant and congruent with background knowledge, then we can say, to the best our knowledge, that we hit upon a causal mechanism that explains a given social phenomenon. But what if one of these conditions fails? A negative result may trigger further research by improving the modelling strategies, or by collecting new data, thus leading to new discoveries that, perhaps, discard background knowledge.

The hypothetico-deductive structure of explanations also allows us to control the goodness of explanation. We can exert (i) a statistical control by measuring, with the coefficient of determination, how much variability is accounted for. We can also exert (ii) an epistemic control, by asking whether results are coherent with background knowledge. (iii) A metaphysical control is also possible, as we have to make sure that there be ontological homogeneity between the variables acting in the mechanism. If such ontological homogeneity is lacking, this would trigger further research for indirect causal paths that would have been previously neglected. A detailed case study illustrating the hypothetico-deductive character of causal-model explanations is discussed in Russo (2008, ch.6.1).

An example: job satisfaction, work values, and job rewards

Let us illustrate with an example. Consider the phenomenon of job satisfaction. A study dated 1977 (Kalleberg 1977) attempted to explain job satisfaction through two variables: work values and job characteristics. Whether or not the results of this study are still valid today will not concern us—the paper still is a good example of the hypothetico-deductive character of causal-models explanations. The importance of correctly explaining job satisfaction goes beyond the cognitive goal of understanding social phenomena—the explanation, as Kalleberg (1977, p.124) says, also has to be useful to improve the work experiences of people. In other words, an action-oriented goal is also at the very basis of causal modelling.

To begin with, this study is not merely an empirical analysis of some data concerning job satisfaction, but it overtly aims at developing a *theory* of this phenomenon. In a footnote Kalleberg (1977, p.126, n.2) points out that his paper is

[9] Simply put, according to Cook and Campbell (1979), *internal validity* establishes whether a relation is causal, or whether from the absence of a relationship between two variables we can infer absence of causality. *External validity*, instead, concerns the possibility of generalising a presumed causal relationship across different times, settings, populations.

> […] an attempt to develop a "theory" of job satisfaction, i.e., a set of generalisations that explains the variation in this phenomenon on the basis of the condition and processes that produce this variation […]. There are two parts to the theory: a "psychological" part which explains the variation in job satisfaction produced by the interplay between work values and job rewards and a "sociological" part, which relates the variation in job satisfaction to factors that affect one's degree of control over the attainment of job rewards.

This quote contains too much important information to be confined in a footnote. Let us disclose it. The first thing worth noting is the explicit explanatory goal of the paper: to pinpoint the factors, the processes, and their interplay—i.e., the mechanism—that underlie job satisfaction and that bring about variations in it. Secondly, the rationale of causality employed is a rationale of variation, not of regularity—the interest is in whether and how variations in key explanatory variables bring about variations in job satisfaction, not whether the latter regularly follows the formers. Thirdly, for the theory to be explanatory, there have to be two parts: one psychological and the other sociological, that is to say Kalleberg needs to model a *mixed mechanism*. I will now go through the paper highlighting (some of) the key elements of the hypothetico-deductive model of explanation discussed above.

The specific modelling strategy chosen here does not come out of nothing, but instead has its origins in previous works in the area. In particular, Kalleberg mentions three approaches. The first attempted to explain variation in job satisfaction merely in terms of personal traits of individuals, but this proved unsuccessful because the relation between job satisfaction and the characteristics of the job—i.e., the sociological part—was neglected. The second approach explained variations in job satisfaction just as a function of differences in the nature of jobs people perform. But this view, argues Kalleberg, had theoretical pitfalls that made it quite useless for a full understanding of the phenomenon, in spite of its practical utility. A third view was that job satisfaction has to be seen as the result of two components: the objective properties of the job and the motives of the individual. However, this approach couldn't establish those links in a systematic way, nor it could provide an adequate conceptualisation of such a mechanism.

These three approaches constitute the background from which the causal hypothesis is formulated and from which the explanatory agenda is set up (Kalleberg 1977, p.126):

> The objectives of this paper are to conceptualize and empirically examine: (1) the way work values and job rewards combine to influence job satisfaction and (2) the factors that determine the extent to which individuals are able to obtain job rewards.

To establish those causal claims will require a thorough discussion of the concepts involved, of the assumptions underlying the model, of the results of empirical testing, of the theoretical plausibility of the causal links, and it will also require a comparison with alternative models. In the paper, the discussion around the concepts involved in the causal model concerns the unitary dimension of 'job satisfaction' and the explanatory role of the variables 'work values' and 'job reward' (see Kalleberg 1977, p.131). An interesting issue relate the possible objection of a correlation between the two dimensions 'value' and 'reward': having defended the choice of the covariate and having provided arguments for their sufficiency in explaining variations in job satisfaction, the next step is to rule out a situation of confounding.[10] Although statistical evidence supports the hypothesis that values and rewards are independent causes of job satisfaction, a theoretical explanation of this results is not provided yet. In other words, Kalleberg has to spell out the mechanism, that is he has to provide "a precise specification of the manner in which particular values and rewards combine to influence

[10] I will not go through the whole arguments provided by Kalleberg. It will suffice to mention that he supports his claims by showing results of performed tests in various tables; in particular, covariate sufficiency is supported by showing the amount of total variation those variables are able to account for.

overall job satisfaction" (Kalleberg 1977, p.132). This is done by regressing job satisfaction simultaneously on values and rewards in an additive and linear model where job satisfaction is a function of rewards and values (plus errors). Testing leads to the following general result Kalleberg (1977, p.133):

> […] the highest levels of job satisfaction will be experienced by those workers with high rewards and low values, while the lowest levels of job satisfaction will be experienced with low rewards and high values.

The question arises as to whether this constitutes a good explanation of the job satisfaction phenomenon. The statistical answer is given by the good value of the coefficient of correlation obtained, and by comparing this coefficient with the one obtained in a different model, that is an alternative model having an interaction effect between values and rewards. Kalleberg argues that this second model would be less appropriate on the following grounds: (i) the first model accounts for more variance than the second, that is it has a better coefficient of determination, and (ii) empirical results of the second model go against the assumptions of the model itself.

Let me skip over the conceptualisation and modelling of the degree of control on job rewards and concentrate instead on the scope and limits of the model. Although Kalleberg makes a good case for the causal hypothesis that work values and rewards have independent effects on job satisfaction, he warns us that the 'job-reward' variable measured *perceptions* of job characteristics. The relationships between actual and perceived job characteristics are indeed an important area to investigate further in order to better understand the source of individual differences in job satisfaction. A second element worth noting is related to the assumption of closure of the system. By way of reminder, causal modelling assumes that it is possible to isolate a mechanism within a larger system, and this mechanism is not subject to external influences. Kallberg's study does not consider other external influences such as the kinds of job characteristics that produce variations in the types of job rewards he analysed. This, he says (Kalleberg 1977, p.140), is a key question for future research. A third consideration on future research consists in investigating what kinds of people have different values towards work, as there is a whole range of social factor that could possibly influence this variable (Kalleberg 1977, p.141). Finally, a full understanding of this phenomenon needs an interdisciplinary analysis, that is it needs modelling a mixed mechanism where social and psychological factors explain job satisfaction.

Kalleberg provides a sound hypothetico-deductive explanation of job satisfaction by following the modelling procedure of causal models. In particular, we are told how the causal hypothesis is formulated, and the explanatory role of the covariates he chooses is supported by statistical and theoretical arguments, thus confirming the causal hypothesis. Also, those results are not written on the stone and further research, as Kalleberg himself indicates, may complete or change the explanation of the job satisfaction phenomenon through a new hypothetico-deductive explanatory procedure.

6. Causal modelling vs. other models of explanation

This last section aims at comparing causal modelling with other models of explanation and at showing why those models are not fully satisfactory in the social sciences and particularly in quantitative social sciences. This comparison will enable us to highlight what causal modelling offers over and above traditional models of explanation.

Contemporary philosophy has been debating explanation for about 60 years now. Salmon (1990a) has brilliantly summarised the first four decades, but much discussion followed since then and, in particular, a novel account—Woodward's manipulationist approach—has been proposed Woodward (2003). I direct the reader to detailed

introductions to explanation (Salmon 1990, Psillos 2002, Woodward 2003)—here I only isolate four main contenders, namely the deductive-nomological model and more generally the covering-law model (Hempel 1965), the statistical-relevance model (Salmon et al. 1971), the causal-mechanical model (Salmon 1984), and the manipulationist model (Woodward 2003). The goal here is not to dismiss those accounts altogether. There is indeed much that can be learnt from them but there are some aspects peculiar to quantitative social science that they are not able, alone, to grasp or to account for.

According to the deductive-nomological model, an explanation is a deductively valid argument where the conclusion, or explanandum, states that the event or phenomenon to be explained occurred. A peculiar characteristic of the D-N model is that the premises of the argument, or explanans, have to contain at least a law. In a D-N explanation there are two types of conditions of adequacy. The first type is logical: (i) the explanation has to be a valid deductive argument, (ii) the explanans has to contain at least a law, and (iii) the explanans has to have empirical content. The second type is an empirical condition: statements in the explanans have to be true. Next to D-N explanations, Hempel also recognised two other types: the deductive-statistical model, where the premises contain at least a statistical law, and the inductive-statistical model, where the explanans confers high probability on the explanandum event.

A first obvious (and well-known) problem concerns laws. First of all, it is a *vexata quaestio* of philosophy what constitutes a law and how we can discern between laws and accidental generalisations. However, this problem becomes even worse in the social sciences because, even if we are prepared to admit that there are laws in the natural sciences, arguably the social sciences do not have laws from which we can deduce the explanandum. A second difficulty in applying the D-N model in social contexts concerns prediction. In D-N explanations the explanandum is an occurred event or phenomenon. However, some causal models are used for forecasting, and the possibility to predict rests on the explanatory and causal power of the factors involved.

However, it might be objected that in spite of some similarity of structure, there is a fundamental difference between the deductive-nomological model and the hypothetico-deductive methodology of causal models. The difference, as Salmon (1990a, Introduction) himself points out, lies in their different goals: in order to explain phenomena, we use hypotheses, laws, or scientific theories that are highly confirmed, whereas in the hypothetico-deductive method the same inferential scheme is used to provide evidence for the hypothesis we want to establish. To this objection I would answer thus: in the hypothetico-deductive model of explanation, the explanatory enterprise becomes a dynamic process that involves hypothesising, deriving the consequences from the hypothesis, and testing the hypothesis against data; in this process, the interplay between establishing generalisations and using those generalisations as background knowledge is fundamental. Thus explanation is not reduced to an inference, but becomes the whole process by which we account for the variability of explanandum by means of the causal mechanism that brings it about.

In this respect the statistical-relevance model developed by Salmon et al. (1971) was a significant step toward approaching explanation and statistical modelling. The motivation behind the statistical-relevance model lies in two problems of the covering-law model in general and of the inductive-statistical model in particular. On the one hand, counterexamples exist showing that not all explanation are arguments, and, on the other, even if the explanans confers to the explanandum a high probability of occurrence, this is not ipso facto a guarantee of the goodness of explanation (see e.g., Salmon (1965) and Jeffrey 1969)). Salmon tried to develop an alternative model of statistical explanation where the principal concept was not that of high probability but that of *statistical relevance*. The main consequence of this shift was that a statistical explanation would now require

two probability values and not only one. In this model of explanation, to explain a fact is to find the narrowest homogenous reference class the fact belongs to. However, the S-R model is too narrow in scope because it essentially applies to contingency tables but not to causal models broadly conceived.

The net advantage of causal modelling over the D-N and S-R explanations is that the generalisations involved need not to be laws. They can be empirical generalisations, weaker than laws but more suitable to the social sciences where we arguably don't have universal and necessary laws. The main flaw of the S-R model, however, as recognised by Salmon himself, is that statistical relevance is not a sufficient condition for causality nor for explanation, and in fact Salmon (1984) developed the causal-mechanical model to solve this problem.

In the causal-mechanical model statistical relevance relations are only the basis upon which a *causal* explanation has to be built. A causal explanation has to appeal to notions such as causal propagation and causal interaction. Those are not explicable in mere statistical terms but require a characterisation in terms of physical notions. In a nutshell, a causal-mechanical explanation aims at tracing the spatio-temporal continuous process in which the cause and the effect occur. The basic concepts of the causal-mechanical model are those of causal process (vs. pseudo-process) and of causal interaction. A causal process is a physical process that is able to transmit marks, namely modifications to the structure of the process that occur as a consequence of a causal interaction between two causal processes. For instance, two billiard balls colliding represent a causal interaction between two causal process, and the mark, i.e. the modification in one or the two trajectories of the balls, persists after the interaction takes place. Instead, the intersection between the shadows of two airplanes is not a causal interaction as no modification persists afterwards. This happens because the two shadows are not causal processes but pseudo-processes.

One might wonder, however, whether the causal-mechanical model, that sees in physical processes and interactions the key to single out causal relations, is applicable in social scenarios too. This is questionable, as I argue in Russo (2008, ch.1), because although complex socio-economic process might well exist, it is not by means of concepts of aleatory causality that we model causal mechanisms in the social domain.

In many ways, the approach the best managed to account for the explanatory import of causal models is the manipulationist or interventionist account developed by Woodward (2003). According to Woodward, causes explain effects because they make effects happen. The bulk of his manipulationist account of explanation rests on the idea that causal and explanatory relationships are potentially exploitable for manipulation and control. More specifically, says Woodward (2003, p.191}, explanation "is a matter of exhibiting systematic patterns of counterfactuals dependence". How and why the counterfactual element comes in will become clearer in a moment. There are two key notions in Woodward's account: causal generalisation and invariance. Causal generalizations are relations between variables and they have the characteristics of being change-relating or variation-relating. Of course, the problem of distinguishing causal from spurious generalisations immediately arises. We could hit upon a change-relating relation that is accidental: for instance, an increased number of storks might be statistically associated with an increased number of births, but arguably there is no causality going on there. Or the change-relating relation might be spurious: yellow fingers might be statistically associated with lung cancer but this is the case because they are effects of a common cause, that is cigarette smoking. So, change-relating relations have to show a certain invariability as prescribed by invariance condition in structural models.

The role of generalisations is worth stressing. In Woodward's account, the role of generalisations goes beyond the role laws played in the D-N model. Here, generalisations

not only (i) show that the explanandum was to be expected, but they also (ii) show how the explanandum would change if initial conditions had changed—this is where the counterfactual element comes in. Generalisations can be used to ask *counterfactual* questions about the conditions under which their explanandum would have been different. This sort of counterfactual information allows us to see that conditions in the explanations are in fact explanatory relevant. So the main advantage of counterfactual explanations over D-N explanations is that whilst the latter can only provide nomic grounds for explaining their explanandum, the former can answer *what-if-things-had-been-different* questions.

But not all counterfactuals will do. Relevant counterfactuals are those that describe the outcome of interventions. To causally explain a phenomenon is to provide information about the factors on which it depends and to exhibit how it depends on those factors. Dependence, and particularly counterfactual dependence, plays a crucial role in Woodward's account. Consider for instance the case of Mr Jones that takes birth control pills and does not get pregnant. In Woodward's account taking birth control pills has no explanatory import in the case of Mr Jones because there is no dependence between this factor and the explanandum (i.e., Mr Jones' not getting pregnant). No intervention on this factor will change whether or not Mr Jones becomes pregnant and therefore this factor lacks any explanatory power. The manipulationist account of explanation has the undisputed merit to tailor the concept of explanation to the actual scientific practice of causal modelling and, particularly, of emphasising the role of *variation*-relating generalisation in answering what-if-things-had-been-different questions. However, this exercise is not pushed far enough.

First, the manipulationist account overlooks the role of background knowledge and this opens the door to non-sense invariant generalisations to be explanatory. In fact, suppose that the relation between the increase of bread prices in England and sea level in Venice were found sufficiently stable, on what grounds could we possibly deny it explanatory power if not on background knowledge? Similarly, in the case of the missed pregnancy of Mr Jones, under Woodward's account we have to appeal to interventions to disclose the non-explanatory role of the factor; however, under the H-D account here developed background knowledge would be, as in the previous case, enough to deny birth control pills any explanatory power. *Second*, Woodward emphasises the explanatory role of counterfactuals that describe the outcome of interventions, but is this always appropriate in the social sciences? There are many factors, such as gender, on which we can't intervene, and yet they play an important causal and explanatory role. *Third* Woodward seems to take for granted that we know what the causes are, and then focuses on what would happen if we intervened on them. However, a causal explanation is also meant to *provide* the causal factors, that is a causal explanation also is the search for causes, not only an answer to what-if-things-had-been-different-questions. Hypothetico-deductive explanations can account for this aspect because they seek to confirm causal hypotheses by including in the model explanatory covariates.

7. Conclusion

Causal modelling aims at explaining social phenomena by modelling causal mechanisms in which relations between variables exhibit a certain structural stability. Causal modelling also makes essential use of the explanatory vocabulary: for the variables, for their role, and for the interpretation of the coefficient of determination. However, in the scientific literature a detailed characterisation of the terms 'explanation' or 'explanatory' is missing. An unsophisticated meaning of 'explanation' is that a phenomenon is explained by a causal model to the extent that we can account for the variation in the response variable by introducing relevant factors as explanatory variables. The coefficient of determination

r^2 quantifies the amount of variation accounted for by the explanatory variables, but, I argued, this statistical answer is insufficient to understand why and how causal models have explanatory power.

I advanced the view that specific causal assumptions contribute to the explanatory import of causal models, notably covariate sufficiency and no-confounding, together with background knowledge. But causal models also participate in the explanation of social phenomena insofar as they model the causal mechanisms that bring them about. Thus causal models can be seen as *models* of explanation having a hypothetico-deductive structure. I emphasised the explanatory role of background knowledge in causal models and argued that alternative models of explanations more often that not, overlook it.

However, this is not tantamount to throwing out the baby with the bath water. The models of explanation here discussed pick out many aspects that contribute to the explanatory import of causal models: drawing consequences from generalisations, evaluating statistical relevance relations, identifying causal mechanisms, answering what-if-things-had-been-different questions. Conceiving of causal modelling as a model of explanation allows us to gather together all these features into a single account. However, this is not tantamount to saying that causal modelling is *the* model of explanation, but that, among various alternatives, this model fits well the case of quantitative social science.

Acknowledgments I wish to thank Robert Franck, Phyllis McKay Illari, Jon Williamson, and Guillaume Wunsch for helpful and stimulating comments on an earlier draft of this paper. Financial support from the F.N.R.S. is also gratefully acknowledged.

References

Bunge M. (2004), "How does it work? The search for explanatory mechanisms", *Philosophy of the Social Sciences*, 34, pp. 182–210.

Caldwell J.C. (1979), "Education as a factor in mortality decline: an examination of Nigerian data", *Population Studies*, 33(3), pp. 395-413.

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

Cartwright N. (2007), *Hunting causes and using them: approaches in philosophy and economics*, Cambridge University Press.

Cook T.D. and Campbell D.T. (1979), *Quasi-Experimentation. Design and Analysis Issues for Field Settings*, Rand MacNally, Chicago.

Dowe P. (2000), *Physical Causation*, Cambridge University Press, Cambridge.

Dupré J. and Cartwright N. (1988), "Probability and Causality: Why Hume and Indeterminism don't Mix", *Noûs*, 22, pp. 521-36.

Eells E. (1991), *Probabilistic Causality*, Cambridge University Press, Cambridge.

Franck R. (2007), "Peut-on accroître le pouvoir explicatif des modèles?", in Leroux A. and Livet P. (eds), *Leçons de philosophie économique*, Economica, Paris.

Hempel C. G. (1965), *Aspects of scientific explanation and other essays*, Free Press, New York.

Jeffrey R. C. (1969), "Statistical explanation vs. statistical inference", in *Essays in Honor of Carl G. Hempel*, N. Rescher (ed), Dordrecht, D. Reidel Publishing Co., pp. 104-113. Reprinted in Salmon et al (1971).

Little D. (2004), "Causal Mechanisms", in Lewis-Beck M.S., Bryman A., Futing Liao T (2004), *Encyclopedia of Social Sciences*, Vol. 1, pp. 100-101.

Lopez-Ríos O., Mompart A. and Wunsch G. (1992), "Système de soins et mortalité régionale: une analyse causale", *European Journal of Population*, 8(4), 363-379.

Kalleberg A.L. (1977), "Work values and job rewards: a theory of job satisfaction", *American Sociological Review*, 42, pp. 124-143.

Mosley W.H. and Chen L.C. (1984), "An analytical framework for the study of child survival in developing countries", *Population and Development Review*, 10, Supplement: Child survival: Strategies for research, 25-45.

Mouchart M., Russo F. and Wunsch G. (2007), "Causality, structural modelling and exogeneity", DP 0709, Institut de statistique, Université catholique de Louvain.

URL= http://www.stat.ucl.ac.be/ISpub/dp/2007/DP0709.pdf

Popper K. (1959), *The logic of scientific discovery*, (translation of *Logik der Forschung*), Hutchinson, London.

Psillos S. (2002), *Causation and Explanation*, Acumen Publishing, Chesham.

Reichenbach H. (1956), *The Direction of Time*, University of California Press.
Russo F. (2006), "The rationale of variation in methodological and evidential pluralism", *Philosophica*,, 77, pp.97-124.
Russo F. (2008), *Causality and causal modelling in the social sciences. Measuring variations*. Springer, New York.
Russo F., Mouchart M., Ghins M. and Wunsch G. (2006), "Causality, structural modelling and exogeneity", Discussion Paper 0601, Institut de Statistique, Université catholique de Louvain, Belgium.
Salmon W.C. (1965), "The status of prior probabilities in statistical explanation", *Philosophy of Science*, 32, pp. 137-142.
Salmon W.C. et al. (1971), *Statistical explanation and statistical relevance*, Pittsburgh University Press.
Salmon W.C. (1984), *Scientific Explanation and the Causal Structure of the World*, Princeton University Press.
Salmon W.C. (1990a), *Four decades of scientific explanation*, University of Minnesota Press.
Salmon W.C. (1990b) "Causal Propensities: Statistical Causality vs. Aleatory Causality, *Topoi*, 9, pp. 95-100.
Salmon W.C. (1998), *Causality and Explanation*, Oxford University Press, Oxford.
Spirtes P., Glymour C. and Scheines R. (1993). *Causation, Prediction, and Search*, New York, N.Y.: Springer-Verlag. 2nd Edition, MIT Press (2001).
Suppes P. (1970), *A probabilistic theory of causality*, North Holland Publishing Company, Amsterdam.
Susser, M. and Susser, E. (1996), "Choosing a future for epidemiology ii: from black box to chinese box and ecoepidemiology", *American Journal of Public Health*, 86, pp. 674–677.
Woodward J. (2003), *Making things happen: a theory of causal explanation*, Oxford University Press.