

DOI: [10.20472/ES.2016.5.3.003](https://doi.org/10.20472/ES.2016.5.3.003)

MEDIATION IN CAUSAL LOG-LINEAR MODELS

GLORIA GHENO

Abstract:

The analysis of the causality is important in many fields of research. To obtain the causal effects in a causal log-linear model I propose a theory, using the odds ratio and the concepts proposed by Pearl, whenever possible. I start calculating the effects in a simple mediation model where I analyze a new interaction effect, which I call cell. After I calculate the effects in a mediation model with two parallel mediators. If they are correlated, Pearl's theory cannot be used and no alternative theory has been proposed so far. To resolve this issue I present a new concept of causality. The mediation models are essential in the marketing, then to it, for a more complex analysis, I apply my causal theory.

Keywords:

Causal effects, customer satisfaction, interaction, log-linear model, mediation, parallel mediators

JEL Classification: C30, C39, M31

Authors:

GLORIA GHENO, Ca'Foscari University of Venice and ECLT, Italy, Email: gloriagheno@libero.it

Citation:

GLORIA GHENO (2016). Mediation in causal log-linear models. International Journal of Economic Sciences, Vol. V(3), pp. 33-49., [10.20472/ES.2016.5.3.003](https://doi.org/10.20472/ES.2016.5.3.003)

Introduction

The causal analysis is important in many fields of research, such as in economics and in the social sciences, because the analyst seeks to understand the mechanisms of the examined phenomena using the relationships among the variables, i.e. the cause-and-effect relationships, where some variables are the causes, others the effects. The term causality defines, therefore, a relationship between an event (the cause) and a second event (the effect), where the latter is interpreted as a consequence of the first event. In general in the analysis of the relations among variables it is, however, important to distinguish between "covariation" and "causation". The study of the "covariation", or interdependence between two variables, which are considered on the same level, is used to measure the strength of the link between them, but not the direction, which assumes an order. A measure of "covariation" is the covariance. The analysis of the causation, or of the cause-and-effect relationships, however, is not "observable". The "covariation" involves, in addition to a theory where the cause-and-effect relationships are known a priori, the directionality of the action and the manifestation of a certain event as a direct and necessary consequence of the occurrence of another event or of a set of events. From this distinction the famous phrase "Correlation does not imply causation" is born. The causal analysis aims to demonstrate the existence of relations, through which one or more cause variables produce one or more response variables, and to measure their intensity. The study of these causal relationships originated mediation analysis which seeks to discover the causal pathways through which the changes are transmitted from the cause to the effect. Mediation analysis originates from the literature of the linear-in-parameters models, and in particular from that of the structural equation model (SEM), going back to Wright (1923) and more recently it is found in the social sciences in the papers written by Baron and Kenny (1986) and by Bollen (1989). In recent decades many authors, including MacKinnon, Hayes and Preacher (MacKinnon et al., 2007; Hayes and Preacher, 2010; Hayes, 2013; Hayes and Preacher, 2014) have continued to study the causal effects in linear contexts or with particular linearity assumptions. Robins (2003) and Pearl (2001, 2009, 2012, 2014) try to find a causal theory applicable to all models but this suffers from problems in specific mediation models in nonlinear cases. Muthén and Asparouhov (Muthén, 2011; Muthén and Asparouhov, 2015) study the theories proposed by Pearl and by Robins applied to SEM with a non-linearity in the variables showing their similarities and their differences.

In mediation analysis the variables can influence in causal way directly, indirectly, or in both ways, other variables. The set of all causal effects which influence a variable is called "total effect". The direct effect is the causal effect of a variable on another variable without intervening variables, while the indirect effect is the causal effect of a variable on another variable, considering it only through the intervention of other variables, called mediators. Hayes (2013) states that a model can have parallel or serial multiple mediators. In the presence of parallel mediators the influencing variable is modeled to influence directly and indirectly the influenced variable, through two or more mediators which do not influence causally each other. This model assumes that the mediators can be correlated or not. In a model with serial mediators these influence causally each

other unlike that with parallel mediators. From the point of view of the identification of the effects these two types of models are similar. Only Preacher and Hayes' causal theory (Hayes and Preacher, 2010; Hayes, 2013) can be applied in a model with serial mediators or with correlated parallel mediators, while that proposed by Pearl (2014) can be applied to them under special linearity hypotheses. Both theories may be, however, applied in a model with uncorrelated parallel mediators. Another type of causal relations is the interaction, which occurs when the effect of a cause variable can depend in some way on the presence or absence of another cause variable and vice versa. A difference between Robins' theory and that proposed by Pearl consists in the calculation of the intensity of the indirect effect in presence of interaction. In his theory Robins states that there is the total indirect effect and that pure, while Pearl considers only the latter. To explain the difference between the two effects I examine a linear model with the mediated variable which is dichotomous (for example 0 if the subject is not subjected to the treatment and 1 if the subject is subjected to the treatment) and I analyze its variation from 0 to 1. The difference between the two effects is due to the interaction between the mediated variable and the mediator (Haferman and Schwartz, 2009). In the total indirect effect the interaction between the mediator and the mediated variable is inserted, while in that pure it is not taken into consideration. In the literature, the interaction effect can be measured on additive or multiplicative scale and in many cases it induces that the effect of a variable on another varies by levels of a third and vice versa. Hayes (2013), unlike Pearl, generally considers a particular version of the interaction, the moderation, in which the effect of a variable on another varies by levels of a third, called moderator, but for which the opposite is not true.

As previously noted the causality is typically studied in the linear models. These, unfortunately, require some features which are not always met by the analyzed data. For example, in most of the cases, the data collected from the surveys are represented by categorical variables. The log-linear models solve with a complementary and compact approach the problem of the multivariate analysis of the categorical variables. They, together with other techniques for the analysis of qualitative data, descend from the linear model, generalizing it with some significant changes (McCullagh and Nelder, 1989). The linear model is a starting point for the development of the regression models and for the analysis of the variance in presence of quantitative data. When a variable of a contingency table is treated as response variable and the others are considered explanatory, the parametric techniques of generalized linear type are very similar to the ordinary regression both in the analysis and in the estimation procedure. The ordinary regression, indeed, is the most suitable model of analysis to consider a response variable as a function of several explanatory variables. A model which is derived from the regression is however also used when a researcher does not want to distinguish the response variables from the explanatory variables.

Figure 1: Simple mediation model with 1 mediator and simple interaction model

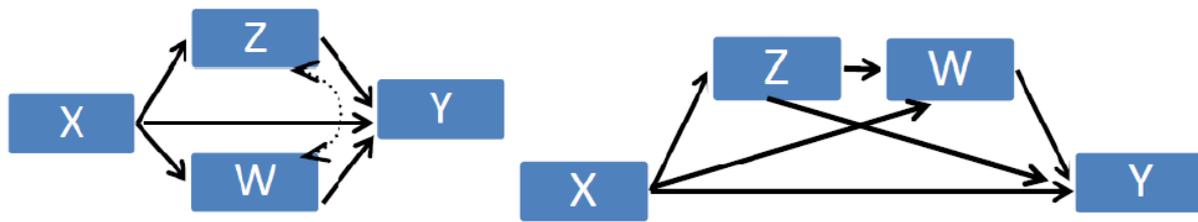
Source: Own path diagrams

In literature the logit models and those log-linear are procedures which are derived from the generalization of the ordinary regression and which perform a quantitative analysis of multivariate frequencies. The logit models assume as criterion variable the logarithm of the odds, i.e. the ratio between the probabilities of the subjects which are in two dichotomous mutually exclusive categories. In the log-linear models the criterion variable is represented directly by the probabilities in the multiplicative version or by the log probabilities in the additive version. The logit models explicitly define a response variable and in this they are opposed to the log-linear models which do not assume, however, causal relationships among the variables. Of course from a logit model it is possible to obtain a log-linear model and vice versa.

In the definition of a logit model it is implicit that no structure among the explanatory variables is taken into account. It is possible, however, to change these logit models incorporating the structure of the explanatory variables, but these models are not like those canonical. To resolve this problem, Goodman (1973) proposes a modeling method to more steps for the response variables, which is called "modified path analysis approach" or causal log-linear model. His procedure involves the construction of a series of marginal tables of incremental complexity which are used to estimate the logit parameters of the various variables.

A problem in the use of the log-linear causal models is, however, the impossibility of calculating all causal effects (Bergsma et al., 2009) and this was, and still is, one of their limits. In this paper, to solve this problem, I propose a new method which calculates the causal effects in a non-linear-in-parameters model, as in that causal log-linear in probabilistic metric, using the odds ratio and a modified version of Pearl's causal theory. My method also solves the problem of the correlated mediators in the non-linear-in-parameters models with the introduction of the new concept of uncorrelated conditional probability. According to Pearl (2014), indeed, if the parallel mediators are correlated, it is possible to calculate the indirect effect only in particular linear models and this is another limitation to the use of nonlinear models.

The remaining part of this paper is organized as follows. Section 2 introduces the log-linear model and its causal version. Section 3 describes my causal theory in a simple mediation model and its version with the introduction of the multiplicative interaction.

Figure 2: parallel multiple mediators model and serial multiple mediators model

Source: Own path diagrams

Section 4 describes my causal theory in a model with uncorrelated and correlated parallel mediators. Section 5 illustrates some applications of my causal theory in marketing.

Log-linear model and causal log-linear model

Before introducing the method to calculate the causal effects, it is useful to underline the passage which leads a log-linear model to becoming a causal log-linear model. It represents a log-linear model in which the variables have a causal role, where, for example, the variable X becomes the cause and the variable Y the effect. Vermunt (1996), too, distinguishes between log-linear models and causal log-linear models. The log-linear model describes the observed frequencies, it measures the strength of the association among variables, but it does not distinguish between response variables and explanatory variables. As mentioned earlier, the causal log-linear model, introduced by Goodman (1973) and also called "modified path analysis approach", is a log-linear model which considers a causal order a priori of the variables. This model is formed by a "recursive" system of logit models in which the variable, which is presented as response in a particular logit equation, can be one of the explanatory variables in one of the consequential equations (Vermunt, 2005). The starting point of Goodman's method is the diagram of the effects proposed by Wright (1921).

Wright proposes a diagram, called path diagram, to represent the relationships among the variables. In the path diagram, the causal direct relationship between two variables is represented by an arrow which goes from the influencing variable to the influenced variable while the correlation between two variables is represented by a double arrow. If two variables are not connected then there is not a direct relationship between them. The first diagram of Figure 1 represents a simple mediation model in which the variable X influences both directly and indirectly the variable Y . The second path diagram of Figure 1 shows a model of interaction, in which the variables X and Z affect directly the variable Y , but in which also their joint effect XZ influences Y . Other examples of path diagrams can be the models with two mediators Z and W illustrated in Figure 2.

To show the difference between the log-linear models and the causal log-linear models, I consider a model with three categorical variables X , Z and Y . In the log-linear model the joint probability of these three variables is defined as:

$$P(X = x, Z = z, Y = y) = \eta \mu^{X=x} \mu^{Z=z} \mu^{Y=y} \mu^{X=x, Y=y} \mu^{X=x, Z=z} \mu^{Z=z, Y=y} \mu^{X=x, Z=z, Y=y} \quad (1)$$

where x , z and y indicate arbitrary categories of the variables X , Z and Y . The interpretation of the parameters of the equation (1) is the following. The parameter η is called overall effect and it is a constant which ensures that the sum of the probabilities is equal to 1 (Heinen, 1996). The parameters $\mu^{X=x}$, $\mu^{Z=z}$ and $\mu^{Y=y}$ are, respectively, the one-variable effects of the categories x , z and y of the variables X , Z and Y . The logarithmic transformation of these parameters indicates the partial skewness of the variables X , Z and Y (Heinen, 1996). The parameters $\mu^{X=x, Y=y}$, $\mu^{X=x, Z=z}$ and $\mu^{Z=z, Y=y}$ are respectively the two-variable effects of the categories x and y of X and Y , of the categories x and z of X and Z and of categories z and y of Z and Y . They can be interpreted as a measure of the strength of the partial statistical relationship between two variables (Heinen, 1996). A practical example may be that of a policy maker who wants to study a state funding for the school abandonment. He might be interested in seeing if the family income affects the continuation of the studies. To understand if there is a disparity in the continuation of the studies depending on the income he will look to the two-variable parameter which measures the relationship between income and school abandonment under the control of a third variable which can be sex. If this parameter is not equal to 1, this means that there is a disparity and therefore he will try to make policies to eliminate it. The parameter $\mu^{X=x, Z=z, Y=y}$ is the three-variable effect of the three categories x , z and y and it measures as odds ratio of two specified variables vary by levels of a third variable (Heinen, 1996). I underline, then, that the three-variable parameter measures the causal multiplicative interaction of X and Z on Y . Recalling the example of the policy maker, the interaction causes the difference due to sex of the relationship between the probability of continuing and that of leaving, both conditioned by two values of income. The interaction, then, shows that the familiar income and the abandonment are associated differently according to sex. In this paper, in particular, I consider the hierarchical log-linear models in which, if a parameter is equal to 1, all the parameters of the upper levels are equal to 1, and then, for example, if $\mu^{X=x, Z=z}$ is equal to 1, also $\mu^{X=x, Z=z, Y=y}$ must be equal to 1. Under this hypothesis, then, if two variables are conditionally independent from the rest, the corresponding two-variable parameter will be equal to 1 (Edwards, 1995).

To simplify I define the variables X , Z and Y binary ($x, z, y = 0, 1$) and because the number of parameters is larger than the number of probabilities ($24 > 8$), the parameters of the equation (1) are not estimable without the addition of constraints. I constrain then the parameters with the dummy code method so that they are identified, that is:

$$\begin{aligned} \mu^{Y=0} &= \mu^{X=0} = \mu^{Z=0} = 1 & (2) \\ \mu^{Z=0, Y=0} &= \mu^{X=0, Y=0} = \mu^{Z=0, Y=1} = \mu^{X=0, Y=1} = \mu^{Z=1, Y=0} = \mu^{X=1, Y=0} = 1 \\ \mu^{X=0, Z=0} &= \mu^{X=0, Z=1} = \mu^{X=1, Z=0} = 1 \\ \mu^{X=0, Z=0, Y=i} &= \mu^{X=0, Z=1, Y=i} = \mu^{X=1, Z=0, Y=i} = \mu^{X=1, Z=1, Y=0} = 1 \quad i = 0, 1 \end{aligned}$$

With these constraints the estimated parameters are reduced to 8 and they become estimable.

The log-linear model does not give a causal order but it only studies the association among the variables. Returning again to the example of the policy maker, the log-linear model studies if there is disparity in the continuation of studies, i.e. it investigates whether abandonment and income are associated, but it does not specify which is the cause and which is the effect. If a researcher wants to study how the income (cause) affects the abandonment (effect) he must proceed to the log-linear causal models. To transform a log-linear model in a causal log-linear model, considering the causal order of the first model of Figure 1, I must assume that the interaction term $\mu^{X=1,Z=1,Y=1}$ is equal to 1, because, if it is present, it introduces the causal multiplicative interaction of X and Z on Y. The presence or absence of this parameter, indeed, determines the presence or absence of the multiplicative interaction which causes that the effect of a variable on another variable depends on a third variable. Following the probability structure proposed by Goodman (1973), the causal model of a simple mediation model can be written as a decomposition of the joint probability in conditional probabilities, i.e. $P(X, Z, Y) = P(Y | Z, X) P(Z | X) P(X)$. If I write the simple mediation model of Figure 1 in causal log-linear terms, I obtain:

$$\begin{aligned}
 P(X = x) &= \frac{\mu_c^{X=x}}{1 + \mu_c^{X=1}} = \eta_c^X \mu_c^{X=x} & (3) \\
 P(Z = z | X = x) &= \frac{\mu_c^{Z=z} \mu_c^{X=x,Z=z}}{1 + \mu_c^{Z=1} \mu_c^{X=x,Z=1}} = \eta_c^{Z|X=x} \mu_c^{Z=z} \mu_c^{X=x,Z=z} \\
 P(Y = y | X = x, Z = z) &= \frac{\mu^{Y=y} \mu^{X=x,Y=y} \mu^{Z=z,Y=y}}{1 + \mu^{Y=1} \mu^{X=x,Y=1} \mu^{Z=z,Y=1}} = \eta^{Y|X=x,Z=z} \mu^{Y=y} \mu^{X=x,Y=y} \mu^{Z=z,Y=y}
 \end{aligned}$$

where c points out the causal log-linear parameters and η is the normalization factor. The causal log-linear parameters are estimated from the conditional probabilities, precisely from the conditional odds, while the log-linear parameters from the joint probability. The parameters of the probability $P(Y = y | X = x, Z = z)$ are the only which are the same in both models. The equations (1) and (2) show an important aspect which differentiates the causal log-linear models from the log-linear models. This differentiation depends on the asymmetric representation which the causal log-linear models give to the relations among the variables. The causal log-linear models include only the parameters which represent the relationships between the cause variables and the effect variables, while a log-linear model can be regarded as a representation of the relationships, bivariate and multivariate, among all variables included in the model.

Causal theory in a mediation model with 1 mediator

In literature the causal log-linear effects are examined in partial way and for this reason, a true causal analysis does not exist. When a researcher considers the simple mediation model of Figure 1, the log-linear literature (Bergsma et al., 2009) calculates, using the odds ratio, the total effect and the direct effect, but it does not study the indirect effect

and this leads to not fully understand the relationships among the variables. The odd is the ratio between the probability of those who are in a category and the probability of those who are not in that category. Its explanation is the possibility that an individual selected in random way is in a predetermined given category rather than in any of the others (Knocke and Burke, 1980). In the example of the policy maker, the odd for the variable abandonment is the ratio between the probability of continuing and the probability of not continuing the studies. It is possible also to calculate the conditional odds: they, given the category of a second variable, define the relationship between the probability of a category and that of the remaining categories of the same variable. In the example of the policy maker the conditional odd is the ratio between the probability of continuing and the probability of not continuing given, for example, the income. The ratio between the conditional odds is called odds ratio. The odds ratio describes the relationship between two binary variables; if the variables are categorical, a transformation in binary variables is required to use it. For example, if I want to analyze the relationship between the categorical variables X and Y with 5 categories, I transform them into binary variables: the transformed variables X and Y are equal to 1, if their original value is 5, 0 otherwise. The relationships analyzed by the odds ratio can be associative or causal (Zhang, 2008). The associative relationships are measured using the effective response, while those causal with the potential response. If the two types of odds ratio are different, this is due to the influence of a third variable called confusion variable (Zhang, 2008; Szumilas, 2010). This confounding variable is causally related to the response variable, but it is not causally linked to the other causal variable (Szumilas, 2010). If the two correlated variables X and Z influence therefore causally the response variable Y, Z becomes a confounding variable of the relationship between X and Y, remembering that the correlation is not a causal relationship. In a simple mediation model without confounding variables, the total effect (TE) and the direct effect used in the log-linear literature (LDE) are given by the following formulas:

$$Or_{x^0, x^1}^{TE} = \frac{P(Y = 1|X = x^1)}{1 - P(Y = 1|X = x^1)} \frac{1 - P(Y = 1|X = x^0)}{P(Y = 1|X = x^0)} \quad (4)$$

$$Or_{x^0, x^1}^{LDE}(Z) = \frac{P(Y = 1|X = x^1, Z = z)}{1 - P(Y = 1|X = x^1, Z = z)} \frac{1 - P(Y = 1|X = x^0, Z = z)}{P(Y = 1|X = x^0, Z = z)} \quad (5)$$

where the subscript (x^0, x^1) indicates that the causal effect measures the effect of the variation of X from x^0 to x^1 . In the next part of my analysis I put $x^0 = 0$ and $x^1 = 1$ by convention. I note, so, that these causal effects coincide with the definitions of total effect and controlled direct effect proposed by Pearl (2001, 2009, 2012, 2014). He, however, never uses the odds ratio to calculate the causal effects, but he prefers the use of the conditional moments. To better understand the log-linear causal models and to apply them to a wide range of case studies, I suggest a causal analysis, using the odds ratio with Pearl's causal theory, when possible, or, otherwise, with his version modified by me. I note that the direct effect is always equal to the causal two-variable parameter $\mu^{X=1, Y=1}$ and that it is independent of the value of the mediator Z. If in a linear-

in-parameters model without interaction the variables X and Z affect the variable Y, but X does not influence Z, the total effect of X on Y is equal to the direct effect of X on Y. This, however, does not occur in a causal log-linear model without interaction: I find, indeed, that, when $\mu^{X=1,Z=1,Y=1}$ is equal to 1, the total effect is not equal to the direct effect, but it involves another effect, which I call cell. The cell effect, therefore, is only present if more variables directly affect the same variable, as in the case in which X and Z influence Y. If there is not the direct effect between X and Y ($\mu^{X=1,Y=1} = 1$) or between Z and Y ($\mu^{Z=1,Y=1} = 1$), the cell effect takes the value 1 and the total effect is equal to the direct effect of Z on Y or of X on Y. The cell formula is the following:

$$\begin{aligned} & Cell_{x^0,x^1}^{effect}(Z) \tag{6} \\ &= \left[\frac{\sum_Z P(Y = 1|X = x^1, Z = z)P(Z = z|X = x^0)}{1 - \sum_Z P(Y = 1|X = x^1, Z = z)P(Z = z|X = x^0)} \frac{1 - P(Y = 1|X = x^0)}{P(Y = 1|X = x^0)} \right] \\ & \quad \left[\frac{P(Y = 1|X = x^1, Z = z)}{1 - P(Y = 1|X = x^1, Z = z)} \frac{1 - P(Y = 1|X = x^0, Z = z)}{P(Y = 1|X = x^0, Z = z)} \right]^{-1} \end{aligned}$$

In this simple mediation model the cell effect, not depending on the value of the mediator Z, can be written as $CellEffect(Z) = CellEffect$, i.e. the cell effect can be interpreted as an effect of constant interaction, but it is important to underline that in a causal log-linear model with multiplicative interaction it is not constant. The cell effect, in general, does not depend on the relationship existing between the variables X and Z, i.e. on the parameter $\mu_c^{X=1,Z=1}$, but only on their joint presence in influencing the response variable Y. To explain more clearly the behavior of the cell effect I use an example in the marketing field. A marketing expert may be interested in customer loyalty analyzing his satisfaction. Customers who purchase a product which they consider good-quality and with a good quality / price ratio (value) will be satisfied and thus they will repurchase more willingly the product. The relation quality-value-satisfaction can be configured as a mediation model where the variable value is the mediator. The cell effect is present when the quality and the product value influence together the satisfaction, but it does not depend on the relationship between the quality and the value, which in this case is causal. The cell effect affects the direct effect of quality on satisfaction and it is born from the only joint presence of quality and value as causes of satisfaction, but it does not depend on the valuation of the product value. The total effect and the direct effect used in the log-linear literature are the odds ratio version of the total effect and of the controlled direct effect provided by Pearl (2009, 2012, 2014). For this reason I propose, in the odds ratio version, his indirect effect

$$\begin{aligned} & Or_{x^0,x^1}^{IE} \tag{7} \\ &= \left[\frac{\sum_Z P(Y = 1|X = x^0, Z = z)P(Z = z|X = x^1)}{1 - \sum_Z P(Y = 1|X = x^0, Z = z)P(Z = z|X = x^1)} \right] \left[\frac{1 - P(Y = 1|X = x^0)}{P(Y = 1|X = x^0)} \right] \end{aligned}$$

and his decomposition of the total effect

$$Or_{x^0,x^1}^{TE} = \underbrace{Or_{x^0,x^1}^{LDE}(Z) Cell_{x^0,x^1}^{effect}(Z)}_{Or_{x^0,x^1}^{NDE}} \frac{1}{Or_{x^1,x^0}^{IE}} \tag{8}$$

to which I add my decomposition of Pearl’s natural direct effect into cell effect and into LD effect. In his theory, he, indeed, has two direct effects: the natural direct effect and the controlled direct effect. The first measures the variation of the variable Y when the variable X varies and the mediator Z remains constant at the value obtained from the initial value of X, while the second measures the variation of Y when X varies and the mediator Z remains fixed to a particular value. Returning again to the example of marketing expert, the natural direct effect measures how the customer satisfaction varies as the product quality varies, considering the product value equal to that achieved by the initial quality. The controlled direct effect, or LD, measures how the customer satisfaction varies as the product quality varies considering fixed the product value.

If the effect of the variable X on the variable Y depends on the value of the mediator Z, I consider a mediation model with a single mediator and with the multiplicative interaction. In this model, then the variable X directly influences Z and the variable Y is influenced directly by the variables X and Z, and by their joint effect, due to the interaction term $\mu^{X=1,Z=1,Y=1}$. My formulas (4), (5), (6), (7) and (8) remain, however, equally valid to calculate the causal effects. The direct effect of X on Y, used in the log-linear literature, becomes a function of the mediator Z, i.e. $\mu^{X=1,Y=1} \mu^{X=1,Z=z,Y=1}$. Consequently, the cell effect also becomes a function of Z. The natural direct, indirect and total effects, however, do not become function of Z. The indirect effect of a model with the multiplicative interaction remains equal to that of a model without the multiplicative interaction, this in perfect agreement with the concept of the indirect effect proposed by Pearl, coinciding with that pure proposed by Robins, as explained in the introduction.

Causal theory in a mediation model with 2 parallel mediators

Very often the variable X, however, affects the variable Y across multiple mediators. Consequently, to arrive at a solution I had to complicate the simple mediation model adding another mediator in parallel way so as to obtain the first model of Figure 2, where the causal effect of the variable X is mediated by two mediators Z and W and where there are no interactions, i.e. $\mu^{X=1,Z=1,Y=1} = \mu^{X=1,W=1,Y=1} = \mu^{Z=1,W=1,Y=1} = \mu^{X=1,Z=1,W=1,Y=1} = 1$. I do not consider the interactions only for simplicity. Following Goodman (1973), the joint probability becomes equal to the product P (Y | Z, W, X) P (Z, W | X) P (X) and the causal log-linear representation of this model results to be

$$\begin{aligned}
 P(X = x) &= \eta_c^X \mu_c^{X=x} \tag{9} \\
 P(Z = z, W = w | X = x) &= \eta_c^{Z|X=x} \mu_c^{Z=z} \mu_c^{W=w} \mu_c^{X=x,Z=z} \mu_c^{X=x,W=w} \mu_c^{Z=z,W=w} \\
 P(Y = y | X = x, Z = z, W = w) &= \eta^{Y|X=x,Z=z,W=w} \mu^{Y=y} \mu^{X=x,Y=y} \mu^{Z=z,Y=y} \mu^{Z=z,Y=y}
 \end{aligned}$$

where the parameter $\mu_c^{Z=z,W=w}$ measures the correlation of the two mediators, or more precisely their association being categorical variables. If the mediators are not correlated ($\mu_c^{Z=z,W=w}=1$), the causal effects can be simply calculated using the odds ratio version of the causal theory proposed by Pearl (2001, 2009, 2012, 2014), who holds separate the analysis of the mediators and who inserts in the natural direct effect that mediated by the mediator not considered. I prefer, however, to unify the indirect effects of the two mediators, examining them together, because this allows me to eliminate the problem, noted by Pearl (2014), of the identification of the natural direct effect in the nonlinear models when the mediators are serial. Returning to the previous example, if the marketing expert studies the relationships among the variables quality-value-satisfaction and future behavior, which measures the customer loyalty, he uses a model with two serial mediators like that of the second diagram of Figure 2, where value and satisfaction are the two mediators Z and W. The mediation of both mediators, if taken together, is measured by the indirect effect and, making so, the natural direct effect becomes identifiable. In the example of the marketing expert, then, I consider the indirect effect of quality on future behavior mediated both by the product value and by the customer satisfaction. The indirect effect, for any separate mediator, may be, however, calculated with the formula (7) in the case of two uncorrelated parallel mediators. In a mediation model with uncorrelated parallel mediators the formulas of the total effect remain equal to those of the model with a single mediator (formulas (4) and (8)) while the direct effect proposed in literature, the cell effect and the indirect effect become, after appropriate modifications, respectively:

$$\begin{aligned} Or_{x^0,x^1}^{LDE}(Z) & \quad (10) \\ &= \frac{P(Y=1|X=x^1,Z=z,W=w) - P(Y=1|X=x^0,Z=z,W=w)}{1 - P(Y=1|X=x^1,Z=z,W=w)} \end{aligned}$$

$$\begin{aligned} Cell_{x^0,x^1}^{effect}(Z) & \quad (11) \\ &= \left[\frac{\sum_{Z,W} P(Y=1|X=x^1,Z=z,W=w)P(Z=z,W=w|X=x^0)}{1 - \sum_{Z,W} P(Y=1|X=x^1,Z=z,W=w)P(Z=z,W=w|X=x^0)} \frac{1 - P(Y=1|X=x^0)}{P(Y=1|X=x^0)} \right] \\ & \quad \left[\frac{P(Y=1|X=x^1,Z=z,W=w)}{1 - P(Y=1|X=x^1,Z=z,W=w)} \frac{1 - P(Y=1|X=x^0,Z=z,W=w)}{P(Y=1|X=x^0,Z=z,W=w)} \right]^{-1} \end{aligned}$$

$$\begin{aligned} Or_{x^0,x^1}^{IE} & \quad (12) \\ &= \left[\frac{\sum_{Z,W} P(Y=1|X=x^0,Z=z,W=w)P(Z=z,W=w|X=x^1)}{1 - \sum_{Z,W} P(Y=1|X=x^0,Z=z,W=w)P(Z=z,W=w|X=x^1)} \right] \\ & \quad \left[\frac{1 - P(Y=1|X=x^0)}{P(Y=1|X=x^0)} \right] \end{aligned}$$

I note that the direct effect proposed in literature remains equal to $\mu^{X=1,Y=1}$ as in the model with a single mediator and without the multiplicative interaction. These equations can also be applied to a model with serial mediators, replacing in them the conditional probability of the mediators Z and W with the product $P(W|X,Z)P(Z|X)$. If I calculate, however, the causal effects with these formulas in a model with correlated mediators, I

would give a causal importance also to the correlation and so I would insert a distortion of the effects due to it. I propose, therefore, a modification to eliminate the role of the correlation so as to remove the distorting effects. Recalling that the conditional probability of the mediators Z and W, given the mediated variable X, is equal to

$$P(Z = z, W = w|X = x) = \eta_c^{Z,W|X=x} \mu_c^{Z=z} \mu_c^{W=w} \mu_c^{X=x,Z=z} \mu_c^{X=x,W=w} \mu_c^{Z=z,W=w} \quad (13)$$

I propose the elimination of parameter $\mu_c^{Z=z,W=w}$, which measures the correlation, so as to consider only the causal relationships and thus the conditional probability P (Z, W | X) becomes

$$\tilde{P}(Z = z, W = w|X = x) = \tilde{\eta}_c^{Z,W|X=x} \mu_c^{Z=z} \mu_c^{W=w} \mu_c^{X=x,Z=z} \mu_c^{X=x,W=w} \quad (14)$$

where the parameters mu are the same as the formula (13). Substituting in the above equations for P (Z, W | X) its new version, called uncorrelated conditional probability, I obtain special formulas so as to be able to calculate the causal effects also in presence of correlated mediators. They are:

$$\tilde{O}r_{x^0,x^1}^{TE} = \frac{\tilde{P}(Y = 1|X = x^1) - \tilde{P}(Y = 1|X = x^0)}{1 - \tilde{P}(Y = 1|X = x^0)} \quad (15)$$

$$\tilde{O}r_{x^0,x^1}^{TE} = \underbrace{O r_{x^0,x^1}^{LDE}(Z) \tilde{C}ell_{x^0,x^1}^{effect}(Z)}_{\tilde{O}r_{x^0,x^1}^{NDE}} \frac{1}{\tilde{O}r_{x^1,x^0}^{IE}} \quad (16)$$

$$\tilde{O}r_{x^0,x^1}^{IE} = \left[\frac{\sum_{Z,W} P(Y = 1|X = x^0, Z = z, W = w) \tilde{P}(Z = z, W = w|X = x^1)}{1 - \sum_{Z,W} P(Y = 1|X = x^0, Z = z, W = w) \tilde{P}(Z = z, W = w|X = x^1)} \right] \left[\frac{1 - \tilde{P}(Y = 1|X = x^0)}{\tilde{P}(Y = 1|X = x^0)} \right] \quad (17)$$

Table 1: causal log-linear parameters

	First dataset	Second dataset	Third dataset
$\mu^{X=1,Z=1,Y=1}$		2.8826*	
$\mu^{X=1,Y=1}$	1.9240**	1.4042	3.9711***
$\mu^{Z=1,Y=1}$	2.4038***	3.5385**	2.9848**
$\mu^{W=1,Y=1}$			9.3261***
$\mu^{Y=1}$	0.4881***	0.2826***	0.0408***
$\mu_c^{X=1,Z=1}$	3.3059***	3.5534***	2.5099***
$\mu_c^{X=1,W=1}$			2.8914**
$\mu_c^{Z=1}$	0.4659***	0.3390***	0.2198***
$\mu_c^{W=1}$			0.0362***
$\mu_c^{X=1}$	1.7132***	1.2278°	0.1832***
Signif. Codes: 0"***" 0.001"***" 0.01"***" 0.05"o" 0.1"***" 1			

Source: Data used as examples in Gheno (2011) and in Gheno (2015)

$$\widetilde{Cell}_{x^0, x^1}^{effect}(Z) = \frac{\left[\frac{\sum_{Z,W} P(Y = 1|X = x^1, Z = z, W = w) \tilde{P}(Z = z, W = w|X = x^0)}{1 - \sum_{Z,W} P(Y = 1|X = x^1, Z = z, W = w) \tilde{P}(Z = z, W = w|X = x^0)} \frac{1 - \tilde{P}(Y = 1|X = x^0)}{\tilde{P}(Y = 1|X = x^0)} \right]}{\left[\frac{P(Y = 1|X = x^1, Z = z, W = w)}{1 - P(Y = 1|X = x^1, Z = z, W = w)} \frac{1 - P(Y = 1|X = x^0, Z = z, W = w)}{P(Y = 1|X = x^0, Z = z, W = w)} \right]^{-1}} \quad (18)$$

where $\tilde{P}(Y|X) = \sum_{Z,W} P(Y|X, Z, W) \tilde{P}(Z, W|X)$. I note that only the equation of the direct effect proposed in the log-linear literature remains unchanged, while all other formulas of the causal effects, having the correlated mediators, logically change.

Numerical studies

In this section I apply my causal theory to empirical data in which the variables are binary with ordered categories (0 = low, 1 = high value). The first two examples consider the relationship among a typical product (in this case Sauris' ham), the satisfaction for its festival and the future behavior of the customers. The third example considers the relationship among the quality of a fast food (in this case Mc Donald), the positive and negative emotions of the costumer and his future behavior. The application of my causal theory in this analysis allows to understand how to stimulate the sales, to increase the flow of tourists in a particular festival or the loyalty of the costumers in commercial business. This analysis is applied in marketing examples, but it can of course be used in many other economic or social sciences sectors.

The first dataset is composed of three dichotomous variables (X measures the interest in Sauris' ham, considering the possibility of buying this product, Z measures the satisfaction about Sauris' Festival, considering the happiness that an individual feels if he thinks he will participate in the event and Y is the future behavior, considering whether an individual will buy the Sauris' ham more often). The parameters estimated by the causal log-linear model are shown in Table 1. The two-variable parameters are all significant, i.e. all different from 1. According to the traditional log-linear literature, the causal two-variable parameters are the causal direct effects. In this case, because all causal two-variable parameters are greater than 1, an increase of the variable X produces an increase of the variable Z, and the same result occurs in the relationships between the variables X and Y, and between the variables Z and Y. Now I calculate the effects using the formulas (5), (6), (7) and (8). The total effect is equal to 2.4008, then an increase of the variable X produces an increase of the variable Y, while the indirect effect is equal to 1.2845, and then an increase of the variable X produces, indirectly, an increase of the variable Y. The cell effect is equal to 0.9741 and, being less than 1, it mitigates the direct effect used in the log-linear literature (LDE). Recalling that the cell effect arises from the presence of two variables influencing Y and that in this case it decreases the direct effect of X on Y, the natural direct effect becomes equal to 1.8741. Consequently, even for the natural direct effect, an increase of X produces an increase of Y. From this analysis, I conclude that if a customer is interested in Sauris' ham, he

will buy this product most often also thanks to the felt satisfaction for his presence at the Festival. In marketing research, this shows that an event linked to the product can increase its sales even if the role of the event is less important than the role of the interest in the product. This result is evident by analyzing the relationship between the indirect effect and the direct effect (indirect effect < direct effect). The joint effect of the interest in the product and of the satisfaction procured by the Festival is due to the mere presence of the two variables which influence a third and it decreases the direct effect used in the log-linear literature, recalling again that the cell effect is less than 1.

Now I consider a second dataset, which yet consists of three dichotomous variables (X measures the interest in Sauris' ham considering the possibility to taste the product, Z measures the satisfaction about Sauris' Festival considering the quality of the products presented during the event and Y is the future behavior considering whether an individual will recommend other people to participate in the Festival). The parameter values are shown in Table 1. It is essential to consider $\mu^{X=1,Y=1}$, although it is not significant, because $\mu^{X=1,Z=1,Y=1}$ is significant. This necessity is due to a feature of the hierarchical log-linear model in which it is required that the n-variable parameters are present when there are the n+1-variable parameters. The total effect is equal to 3.1886, then an increase of X produces an increase of Y. The indirect effect is equal to 1.4493 then an increase of the variable X produces, indirectly, an increase of the variable Y. Now I consider the effect of the multiplicative interaction, with parameter greater than 1, which influences the LD effect and the cell effect. The cell effect with Z = 1 is 0.42705, mitigating thus the LD effect, while with Z = 0 it is 1.2310, increasing thus the LD effect. When the mediator Z is high (Z = 1), the LD effect is equal to 4.0477 (i.e. $\mu^{X=1,Y=1}\mu^{X=1,Z=1,Y=1}$), while when the mediator Z is low (Z = 0), it is equal to 1.4042 (i.e. $\mu^{X=1,Y=1}$). For each value of satisfaction, then, the effect of total interaction (effect cell (Z) x $\mu^{X=1,Z=z,Y=1}$) is always equal to 1.2310, and consequently the natural direct effect is always equal to 1.7286. From the results of this analysis, I conclude that if a customer is interested in Sauris' ham, then he will advise other people to participate in the event more often, thanks to the quality of the products presented and to the overall combined effect of interest and of satisfaction.

The third dataset is instead composed of four dichotomous variables (X measures McDonald atmosphere, considering the music in the fast food, Z measures the positive emotions, W measures the absence of negative emotions and Y is the desire of returning of the costumer). The interactions, being not significant, are not considered. The parameter, which determines the presence of the correlation between Z and W, is significant ($\mu_c^{Z=1,W=1} = 4.1166$, p-value = 0) and for this reason I must to apply the formulas (15), (16), (17) and (18). The remaining parameters are significant and they are shown in Table 1. In this case, since the causal two-variable parameters are all greater than 1, an increase of the influencing variables produces an increase of the influenced variables, according to the traditional log-linear literature. The total effect is equal to 5.1348, the indirect effect is equal to 1.5664, the cell effect is equal to 0.9035 and then the natural direct effect is less than the direct effect of the log-linear literature

(3.5879 < 3.9711). From the results of this analysis I conclude that a good atmosphere of the fast food (quality of the commercial business) positively influences, both directly and indirectly, the return of the customers.

Conclusions

When a researcher analyzes the data, he is interested in understanding the mechanisms which regulate the changes of the variables, and to know them, he uses the causal effects. If he wants to use the log-linear models to study the data, unfortunately, he has not available an exhaustive causal theory, but only few comments on some papers in which the odds ratios are used. The causal theory, indeed, focused mainly on the linear models or on models in which some kind of linearity for the mean of the variables is assumed. In a causal log-linear model, however, there is no assumption of linearity for the mean of the variables. Among the causal authors Pearl (2001, 2009, 2012, 2014) is one of the few who deals with causality applicable to any type of model. To overcome these various limitations and problems I used the causal concepts outlined by Pearl (2001, 2009, 2012, 2014) but formulating a new causal theory to calculate the causal effects in the log-linear models and using the odds ratio so that the parameters have the same interpretation of the log-linear literature. During the development of my theory I found a new effect which I called cell effect. It can be interpreted as an interaction effect which occurs each time that I consider two variables which directly affect a third. Another limitation in the log-linear model is related to the calculation of the indirect effect when the parallel mediators are correlated. To solve this problem I propose a new concept, which I call the uncorrelated conditional probability, which, used in the formulas of the model with uncorrelated mediators, allows to calculate the causal effects eliminating the problem of the correlation. In conclusion my causal theory, applied to the causal log-linear models, allows the calculation of the causal effects for each type of path diagram, and then, with its use, it is always possible to measure their intensity.

References

- BARON, R & KENNY, D. (1986). The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51, pp. 1173-1182
- BERGSMA, W. P., CROON, M. A., & HAGENAARS, J. A. (2009). *Marginal models for dependent, clustered, and longitudinal categorical data*. New York: Springer.
- BOLLEN, K. (1989). *Structural equation with latent variables*. New York: John Wiley
- EDWARDS, D. (1995). *Introduction to Graphical Modelling*. New York: Springer
- GHENO, G. (2011). *Modelli strutturali con variabili categoriali: un approccio a classi latenti*. Master Thesis, University of Padua.
- GHENO, G. (2015). *Structural equation models with interacting mediators: theory and empirical results*. Phd Thesis, University of Padua.

- GOODMAN, L. A. (1973). The analysis of multidimensional contingency tables when some variables are posterior to others: a modified path analysis approach. *Biometrika* 60(1), pp.179–192.
- HAFERMAN, D. M. & SCHWARTZ, S. (2009). Opening the Black Box: a motivation for the assessment of mediation. *International Journal of Epidemiology* 38, pp 838-845
- HAYES, A. F. (2013). *Introduction to mediation, moderation and conditional process analysis: a regression-based approach*, New York: The Guilford press.
- HAYES, A. F., & PREACHER, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research* 45, pp 627-660
- HAYES, A. F., & PREACHER, K. J. (2014). Statistical mediation analysis with multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology* 67, pp 451-470
- HEINEN, T. (1996). *Latent class and discrete latent trait model: similarities and differences*, Thousand Oaks: Sage Publications.
- KNOKE, D. & BURKE, P. J. (1980): *Log-linear models*, Beverly Hills: Sage Publications
- MACKINNON, D., FAIRCHILD, A., & FRITZ, M. (2007). Mediation analysis. *Annual Review of Psychology* 58, pp 593-614
- MCCULLAGH, P. & NELDER, J.A. (1989). *Generalized Linear Models, Second Edition*. London: Chapman and Hall.
- MUTHÉN, B. (2011). *Applications of causally defined direct and indirect effects in mediation analysis using SEM in MPLUS*. Technical Report, Los Angeles: Muthén & Muthén
- MUTHÉN, B., & ASPAROUHOV, T. (2015). Causal effects in mediation modeling: an introduction with applications to latent variables. *Structural Equation Modeling: A Multidisciplinary Journal* 22(1), pp 12-23
- PEARL, J. (2001). Direct and indirect effects. In J. Breese and D. Koller ed. *Proceedings of seventeenth conference on uncertainty and artificial intelligence*, San Francisco: Morgan Kaufman, pp. 411-420
- PEARL, J. (2009). Causal inference in statistics: An overview. *Statistics Surveys* 3, pp. 96–146.
- PEARL, J. (2012). The mediation formula: a guide to the assessment of causal pathways in nonlinear models. In C. Berzuini, P. Dawid, and L. Bernardinelli ed. *Causality: Statistical Perspectives and Applications*, Chichester: John Wiley & Sons, pp. 151–179.
- PEARL, J. (2014). Interpretation and identification of causal mediation. *Psychological Methods* 19, pp. 459-481
- ROBINS, J. M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In P. Green, N. L. Hjort and S. Richardson ed. *Highly structured stochastic system*, New York: Oxford University Press, pp. 70-81.
- SZUMILAS, M. (2010). Explaining odds ratios. *Journal of the Canadian Academy of Child and Adolescent Psychiatry* 19(3), pp. 227–229.
- VERMUNT, J. K. (1996). Causal log-linear modeling with latent variables and missing data. In U. Engel and J. Reinecke ed. *Analysis of Change: Advanced Techniques in Panel Data Analysis*, Berlin/New York: Walter de Gruyter, pp. 35–60.
- VERMUNT, J. K. (2005). Log-linear models. In B. Everitt and D. Howell ed. *Encyclopedia of Statistics in Behavioral Science*, Chichester: Wiley, pp. 1082–1093
- WRIGHT, S. (1921). Correlation and causation. *Journal of agricultural research* 20(7), pp. 557–585

ZHANG, Z. (2008). Estimating a marginal causal odds ratio subject to confounding. *Communications in Statistics-Theory and Methods* 38(3), pp. 309–321.