# RESEARCH NOTE: GRAPHICAL MODELS OF CAUSATION

Paul Hünermund*

This version: 29 November 2017
First version: 25 October 2017

The computer science and artificial intelligence literature provides powerful tools for causal inference with observational data based on graphical models of causation. A paper by Durand and Vaara (2009), which introduced causal graph modeling to the strategy literature, has been criticized by Ellsaesser *et al.* (2014) in *Strategic Management Journal*. In this paper I develop a counter-critique of Ellsaesser *et al.* and show that the objections they raise are unjustified. I then proceed to illustrate the usefulness of causal graphs for theory building and empirical research in strategic management.

**Key words**: causal inference; causality; causal graph modeling; directed acyclic graphs; causal structure learning
**JEL classification**: C10, M20

## INTRODUCTION

In the last three decades, influential contributions to the understanding and operationalization of causality have been made in the computer science and artificial intelligence literature (Spirtes *et al.*, 2001; Pearl, 2009). The proposed approaches combine mathematical graph theory with probability concepts from statistics to

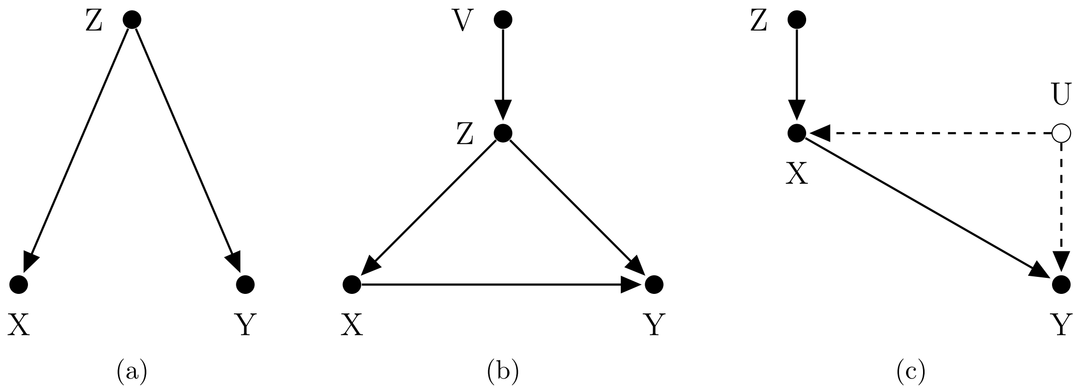
*Maastricht University, School of Business and Economics. Tongersestraat 53, 6211 LM Maastricht, Netherlands. E-mail: p.hunermund@maastrichtuniversity.nl

establish the epistemological foundation of a *theory of causation* and to develop effective tools for empirical causal research. Graphical representations of causal models (*causal graphs* hereafter) encode assumptions about cause-and-effect structures underlying a set of empirical observations. They give rise to conditional independence relationships, which are at the heart of causal inference. Causal graphs are close in spirit to path diagrams (Wright, 1921, 1934) commonly used in traditional linear structural equation modeling (SEM). They relax the assumption of linearity in path diagrams and, as such, provide a non-linear generalization of SEM[1] (Heckman, 2005). Informative criteria and powerful algorithms based on graphical representations have been developed to discern causal effects from mere statistical association in observational data (i.e., in the absence of experimental manipulation).

Causal graphs have attracted attention in various disciplines, such as economics (Neuberg, 2003; Heckman and Pinto, 2013; Spiegler, 2016), finance (Yang and Zhou, 2016), accounting (Gow *et al.*, 2016), sociology, and epidemiology (see Elwert, 2013, for references from the last two disciplines). Durand and Vaara (2009) introduced them to the field of strategic management and proposed graphical models as an apparatus for counterfactual reasoning, which—according to the authors—is still poorly understood in the strategy field. Since then, however, diffusion of this new technique has been slow. The reason for it might be a paper by Ellsaesser *et al.* (2014) in *Strategic Management Journal*, which questions the applicability of causal graphs in empirical management research and recommends the use of *vector space modeling* as an alternative. In this paper I develop a counter-critique

[1] Although the literature on causality in computer science was initially developed independently from the SEM literature, both strains have since found together and cross-fertilized each other (Bollen and Pearl, 2013; Kline, 2015).

Figure 1: Examples of causal graphs




of Ellsaesser *et al.* and show that the objections they raise are unjustified. For the sake of exposition the following section provides a short introduction to graphical models of causation and illustrates their use with an example from management theory. I then proceed to demonstrate the usefulness of causal graphs for theory building and empirical research in strategic management.

## CAUSAL GRAPH MODELING

A graph is a mathematical object consisting of a set of *nodes*, representing the variables in a model (e.g., $X$, $Y$, and $Z$ in Figure 1a), and a set of *edges* connecting the nodes. Edges are directed, $Z \rightarrow X$, and arrowheads denote cause-and-effect relationships between a *parent* and a *child* node. A graph is acyclic if there are no directed causal paths—i.e., chains of arrows—from a variable pointing back to itself. This notion precludes feedback loops in which a variable would be a cause of itself.[2] Nodes representing unobserved variables are indicated by hollow circles and the edges they emit remain dashed (see $U$ in Figure 1c). This signifies that,

[2]Acyclicity refers to what is also known as a *recursive* system in the SEM literature.

although a variable is integral part of a larger causal model, it is not measured by the analyst.

*Directed acyclic graphs* (DAG) are central to graphical models of causation. Every DAG is associated with a *structural causal model* (SCM), that represents variables in their functional form. E.g., the SCM for the graph in Figure 1b is given by

$$V = f_1(\varepsilon_1), \tag{1}$$

$$Z = f_2(V, \varepsilon_2), \tag{2}$$

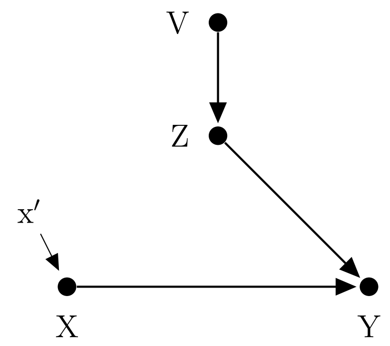
$$X = f_3(Z, \varepsilon_3), \tag{3}$$

$$Y = f_4(X, Z, \varepsilon_4). \tag{4}$$

The $\varepsilon_i$'s are mutually independent random variables representing errors, or omitted factors (Pearl, 2009, ch. 1.4.1), that induce a joint probability distribution of the model. As the $f_i$'s remain unspecified, DAGs are fully non-parametric.

Note that conditional on its immediate parents a variable is independent of its non-descendants (i.e., its parents' ancestors). For example, by conditioning on $Z$ in graph 1b, $X$ only depends stochastically on $\varepsilon_3$ and is therefore independent of $V$. This so-called *Markov condition* (Pearl, 2009, ch. 1.2) forms the backbone of causal inference in graphical models. It implies that the causal effect of $X$ on $Y$ is empirically identified in the causal model represented by graph 1b, and can therefore be estimated. Conditioning on $Z$ closes the only *backdoor path* (Pearl, 1993)—i.e., an alternative causal connection producing and association between $X$ and $Y$—remaining in the model. Consequently, since $X$ becomes independent

Figure 2: Postintervention graph of Figure 1b




of any other variable affecting $Y$, a confounding bias stemming from unaccounted influences can be excluded.

In line with the notion of controlled experiments, causal effects are defined in terms of interventions. These constitute deliberate manipulations of variables in a model, whose values are otherwise chosen by nature.[3] Intervening on $X$ in graph 1b means to fix $X$ at a particular value, $x'$—an operation which is denoted by $do(X = x')$. Algebraically, equation 3, the natural data generating process of $X$, is substituted by 3': $X = x'$. Since all influences that would usually determine $X$ are eliminated by the intervention, edges pointing to $X$ are omitted in the post-intervention graph (see Figure 2).

The causal effect of $X$ on $Y$ in is given by the post-intervention probability distribution $P(Y|do(X))$ in the modified graph. An amazing result, established by the work of Judea Pearl, is that, although interventions have to remain a purely theoretical concept in ex-post data analysis, if randomized experiments are infeasible, the post-intervention distribution in graph 2 can be expressed solely in terms of observable quantities (see Pearl *et al.*, 2016, ch. 3.2, for a proof):

[3] Note, however, that in the context of ex-post data analysis, interventions are only considered to be a mental, or hypothetical, operation.

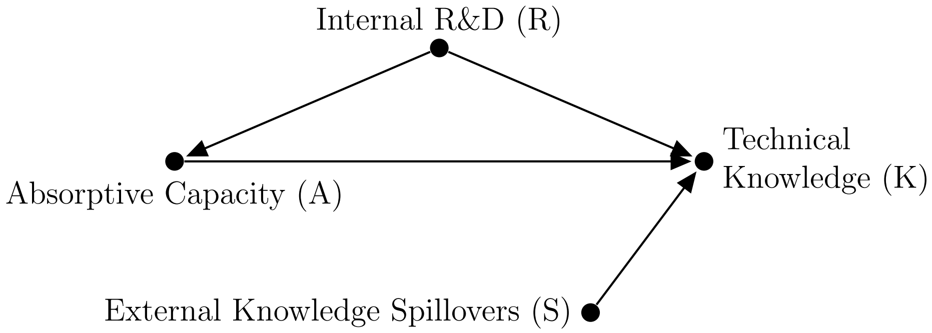
$$P(Y|do(X)) = \sum_{Z} P(Y|X,Z)P(Z).$$

Thus, estimating the causal effect of $X$ on $Y$ with the help of ex-post observed data involves determining the conditional probability of $Y$ given $X$ in every strata of $Z$, and then averaging over $Z$. This result provides the theoretical justification for why conditioning on (also denoted as "adjusting for" or "controlling for") $Z$ is sufficient in graph 1b.

For a concrete example from management theory, consider the model of absorptive capacity, which Cohen and Levinthal (1990) illustrated graphically in their seminal article (see Figure 3). The extent to which a firm benefits from external knowledge, through spillovers from competitors or extra-industry sources such as universities, depends on its absorptive capacity. At the same time, because absorptive capacity *"is often generated as a byproduct of [a firm's own] R&D"* (Cohen and Levinthal, 1990, p. 138), the theory postulates a benefit of internal R&D beyond its immediate (direct) effect on a firm's knowledge stock.

In order to empirically assess whether absorptive capacity indeed alters the extent to which a firm can benefit from external knowledge spillovers, i.e., whether there is a combined effect of absorptive capacity and external knowledge spillovers on a firm's technical knowledge, one needs to close the back-door path going through internal R&D ($A \leftarrow R \rightarrow K$). Since R&D both affects a firm's absorptive capacity and technical knowledge, it acts as a confounder and introduces estimation bias if not properly dealt with. It is thus necessary to adjust for the level $R$ in order to identify the post-intervention distribution of $K$ (with respect to $A$ and $S$)

Figure 3: A DAG of absorptive capacity and technical knowledge




*Notes:* Causal graph adapted from Figure 2 in Cohen and Levinthal (1990). In the original figure, the arrow originating from $A$ points to the arrow connecting $S$ and $K$ instead of the node $K$. Such an "arrow-on-arrow" representation (Weinberg, 2007) indicates effect modification (or moderation). In a DAG direct effect modification can be expressed as both the modifying and modified variable emitting arrows to the outcome variable (VanderWeele and Robins, 2007).

$$P(K|do(S,A)) = \sum_R P(K|S,A,R)P(R) \tag{5}$$

From the right-hand side expression, which depends solely on observable quantities, the average causal effect (ACE; Pearl, 2009) of a change of $S$ from $s$ to $s'$ for $A$ held fixed at $a$

$$E[K|do(S=s', A=a)] - E[K|do(S=s, A=a)]$$

can be computed. If the ACE of $S$ is different for different levels of $A$, then absorptive capacity moderates the causal relationship between external knowledge spillovers and a firm's technical knowledge, as originally proposed by Cohen and Levinthal.[4]

[4]Due to space limitations this introduction to causal graphs has to remain short and incomplete. Pearl *et al.* (2016) and Morgan and Winship (2007) are good introductory-level texts. In-depth coverage is provided by Pearl (2009) and Spirtes *et al.* (2001).

## CRITIQUE OF ELSAESSER, TSANG, AND RUGE

Ellsaesser *et al.* (2014) question the applicability of causal graphs in situations typically encountered in strategic management research and put forward two points of criticism. Their first example is concerned with the problem of missing information. Consider the graph in Figure 1c (analogous to Figure 3 in Ellsaesser *et al.*), in which $U$ acts as an unobserved confounder between $X$ and $Y$. Durand and Vaara (2009) introduce the so-called "front-door" criterion (Pearl, 2009, ch. 3.3.2) as one of three strategies for causal effect identification. In graph 1c front-door adjustment would involve to condition on $X$ in order to identify the causal effect of $Z$ on $Y$. This, however, is rendered impossible by the presence of the unobserved $U$, as Ellsaesser *et al.* correctly note.[5]

It is true that missing data pose serious challenges for the identification of causal effects in empirical work. However, this problem is hardly unique to causal graph modeling and therefore does not discredit graphical models in themselves. On the contrary, graphs make explicit in which cases unobservables cause problems for identification and in which they do not. Careful inspection of graph 1c, for example, leads to the conclusion that invoking the front-door criterion is unnecessary for identification. Because $Z$ is exogenous, with no parents affecting it, the post-intervention distribution $P(Y|do(Z))$ is identified without the need to control for any other variable. Relying more thoroughly on causal graph methodology would have prevented Ellsaesser *et al.* from overlooking this much simpler source of identification in their model.

In a second example, Ellsaesser *et al.* question how well the Markov condition,

[5] Technically, $X$ acts as a *collider*, which would open up the confounding path $Z \rightarrow X \dashleftarrow U \dashrightarrow Y$ if it were conditioned on (see Pearl, 2009, ch. 1.2.3, for details).

which gives rise to conditional independence relations and is central to graphical models of causation, represents reality. They discuss a situation as in graph 1a, where $Z$ stands for the installation of a new machine in a firm's production process. $Z$ is assumed to causally affect both the defect rate of output ($X$) and the chance of machine breakdown ($Y$). It is easy to see that, given the assumptions encoded in the graph, $X$ and $Y$ are independent of each other, conditional on $Z$. Consider the SCM representation of graph 1a,

$$X = f_1(Z, \varepsilon_1), \quad Y = f_2(Z, \varepsilon_2).$$

Conditional on $Z$, both $X$ and $Y$ depend stochastically only on the independent errors $\varepsilon_1$ and $\varepsilon_2$ and are therefore independent themselves.

Ellsaesser *et al.* do not find this convincing: *"since both types of events result from a common cause, it is surely reasonable to expect that their occurrences are correlated"* (p. 1545). Indeed, over-simplified models such as 1a can only provide an unsatisfactory representation of the complex mechanisms at play in many real-world situations. For instance, it might be more realistic to introduce an unobserved confounder $W$ in the model ($X \dashleftarrow W \dashrightarrow Y$) that produces a correlation between $X$ and $Y$, even conditional on $Z$.[6] However, once again this argument does not identify a weakness of causal graph modeling in itself. The converse is true: graphical models make explicit the structural assumptions (e.g., whether a confounder is absent or not) that are necessary to justify conditional independence theoretically.

[6] An equivalent notation is to introduce a *bidirected* edge, $X \dashleftarrow\dashrightarrow Y$, denoting a not further specified common cause of $X$ and $Y$ (Pearl, 2009, p. 12) that produces a correlation in the associated error terms. Graphical causal models with bidirected edges are called *semi-Markovian.*

As conditional independence (also known as *ignorability*) is routinely invoked in causal effect estimation (Imbens, 2004; Angrist and Pischke, 2009), and, e.g., constitutes an integral part of instrumental variable techniques (Imbens and Angrist, 1994), simply dismissing it as implausible per se appears to be inadmissible. This would imply that the identification of any causal mechanism in strategic management research based on observational data is a hopeless endeavor. Whether Ellsaesser *et al.* truly intended such a harsh verdict remains unclear; in any case, they provide no justification for it.

## ADVANTAGES OF CAUSAL GRAPHS

From the previous discussion it becomes clear that the criticism of causal graph modeling put forward by Ellsaesser *et al.* (2014) misses the point. Instead, Durand and Vaara (2009) are right when they remark that *"the combination of causal graphs and counterfactual testing and evaluation provides powerful tests of causal relationships"* (p. 1255). The thrust of graphical methods becomes particularly apparent in three distinct domains, (1) theory building, (2) empirical testing, and (3) causal learning, which I will discuss individually below.

### A convenient way to formalize theory

Theory in strategic management research—and other social sciences—is predominantly communicated verbally (Loch and Wu, 2007). This type of theorizing requires a researcher to specify causal relationships between the variables under study (Gow *et al.*, 2016). In the example of Figure 3, these were internal R&D, external knowledge spillovers, absorptive capacity and a firm's technical knowledge. At the same time, however, the exact functional form of the causal links usually remains unspecified. This is often regarded as an advantage of verbal theorizing because the contemplated theoretical mechanisms are thought to apply independently of a specific quantitative formulation; in the sense that a statement "output, $Y$, is created by employing the inputs capital, $K$ and labor, $L$" is more general than "$Y = L^{0.7} \cdot K^{0.3}$".

Causal graphs do exactly that. They encode qualitative causal assumptions but do not necessitate explicit functional forms, which is their main advantage over traditional linear SEM (Pearl, 2012). Consequently, they are well suited to complement the verbal exposition of theory. In fact, management scholars frequently make use of directed acyclic graphs for visual communication of their theories (see, for example, Powell *et al.*, 2006, and Donaldson *et al.*, 2013), without necessarily noticing their empirical content. To fill this gap, I will discuss the close link that causal graphs provide between theory formulation and estimation in the following.

## A unified framework for empirical research

Causal graphs bridge the gap between theory and empirics. Once theory has been represented graphically, a set of inference rules called *do-calculus* (Pearl, 1995, 2009) can be applied to identify the causal effect of an intervention in one variable of the model on other variables of interest. In the example of Figure 1b identification required to compute effects separately for all values of $Z$ and then to average over the distribution of $Z$. With the help of graphical reasoning it was established that adjusting for $Z$ is sufficient, and the variable $V$ can thus be ignored in the estimation process. Do-calculus, however, is applicable in far more complex situations and for larger graphs than 1b. Only in these cases the full potential of graphical models becomes apparent.[7]

Perhaps the biggest strength of do-calculus is its *completeness* property, which means that if it fails to return a queried causal effect, no such effect is identifiable under the given assumptions (Pearl, 2013). Powerful algorithms such as the one developed by Shpitser and Pearl (2006) then turn an otherwise tedious identification task into a straightforward exercise. They are implemented in easy-to-use software packages such as `DAGitty`[8] (Textor *et al.*, 2011) or `causaleffect`[9] for `R`. Researchers only need to submit the description of a graphical model to the software, which then returns causal effect expressions similar to the one in equation 5.

Obviously, it is not always possible to identify causal effects simply from observational data. In cases when unobservables prohibit identification, causal graphs provide useful information about variables that might be suitable candidates for randomization in a controlled experiment, in order to answer the research question at hand. Another useful extension of graphical models is that, although the theory is fully non-parametric, once sufficient sets for covariate adjustment are found, it is easy to introduce further parametric assumptions such as linearity or non-linear link functions (Nelder and Wedderburn, 1972). This can be advisable in situations where non-parametric methods are underpowered because of limited data (Pearl,

[7]If the causal relationships between variables under study is complex, researchers often try to control for as many covariates as possible. However, in many situations this strategy is either unnecessary or downright inadmissible. An example for the latter is given by graph 1c in which adjusting for $X$ would introduce estimation bias because of the unobserved confounder $U$. The graphical representation, by contrast, uncovers that no adjustment is needed to identify $P(Y|do(Z))$.

[8]`http://www.dagitty.net`

[9]`http://cran.r-project.org/web/packages/causaleffect/index.html`

2012).[10] Lastly, causal graphs naturally incorporate tools such as mediation analysis (Imai *et al.*, 2010, 2011; Pearl, 2014), which has many important applications in management research (Vancouver and Carlson, 2015).

#### Causal structure learning in a world of big data

Another very useful feature of graphical methods is the possibility to learn the causal structure underlying a set of variables from data. The traditional research paradigm requires scholars to specify a causal model from substantive theoretical knowledge upfront, which can later be taken to the data (Gow *et al.*, 2016). Likewise, in the examples discussed so far, graphs have been determined by theoretical arguments (for instance, in the absorptive capacity example of Figure 3). As argued before, these graphs encode conditional independence assumptions, such as $X \perp\!\!\!\perp Y \mid Z$ in 1a, which can then be invoked to estimate causal effects. *Causal structure learning* (also denoted as *causal search* or *structure discovery*), by contrast, turns this argument on its head. From a set of testable conditional (in-)dependence relationships between variables in a data set, one can draw conclusions on the class of graphs that is compatible with them (Spirtes *et al.*, 2001; Pearl, 2009). In that way the data reveal information about their own data generating process.

The intuition behind the approach is easily understood (see, e.g., Alemi *et al.*, 2016). Consider a set of three variables, $A$, $B$, and $C$. Under the assumption of acyclicity, they can form three possible causal structures relating to each other: (1)

[10] Invoking functional form assumptions might also help to increase identifying power (Lewbel, 2016). An example for this is the well-known instrumental variable estimator, which actually is non-parametrically unidentified (Pearl, 2009, p. 90) and only becomes identified by assuming monotonicity in the first stage (Imbens and Angrist, 1994).

a causal chain, $A \to B \to C$, (2) a common effect, $A \to B \leftarrow C$, or (3) a common cause, $A \leftarrow B \to C$. Now, if we encounter a situation in which two initially independent variables, $A$ and $C$, become stochastically dependent conditional on another variable $B$, this must be due to a common cause structure.[11] Although, it is not possible to distinguish between the two other causal structures based on empirical observation alone, repeated conditional independence tests can reveal information about common cause configurations for all possible triplets in a data set.[12] In that way, often enough information about the underlying graph can be learned, such that meaningful inference becomes possible.[13]

Causal structure learning has been applied, among others, by Yang and Zhou (2013) to understand the transmission of credit risk between international financial institutions during the 2007-2008 financial crisis. The approach is particularly promising when data sets are large and contain a rich set of variables. It can then complement other machine learning techniques that are traditionally focused on prediction rather than causation (Varian, 2014). The package `pcalg` for `R` (Kalisch *et al.*, 2012), for example, provides extensive functionality to combine the power of structure learning with subsequent causal effect estimation[14].

[11] Formally, the requirement is $A \perp\!\!\!\perp C \mid \emptyset$ and $A \not\perp\!\!\!\perp C \mid B$. Again, $B$ acts as a collider in a common cause structure, which, once conditioned on, renders $A$ and $C$ dependent by opening up the path between them.

[12] This logic of "local" tests is very different from approaches in traditional linear SEM, where predominantly global goodness of fit measures are used to discriminate between models (Kline, 2015).

[13] Stated more formally, causal structure learning can recover DAGs from data up to a certain equivalence class (Pearl, 2009, ch. 2.5). Whether this class offers enough substantive insights in a particular context needs to be decided on a case-by-case basis.

[14] Also see the Tetrad Project by Glymour, Scheines, Spirtes, and Ramsey: `http://www.phil.cmu.edu/tetrad`

# CONCLUSION

Causal graph modeling is a promising tool for research in strategic management. Graph theory provides a rigorous theoretical framework for thinking about causality and counterfactuals (Durand and Vaara, 2009; Donaldson *et al.*, 2013). It also establishes a close connection to statistical methodology and the mechanics of estimation, which is an often overlooked benefit of the approach. Moreover, based on graphical models, causal structure learning bears the promise to uncover cause-and-effect relationships from data. This offers an alternative to the exclusively theory-driven approach to causation that is dominant today (Gow *et al.*, 2016).

Nevertheless, eight years after Durand and Vaara (2009) introduced causal graphs to the strategic management community, the technique has still not been adopted widely. I suspect that this has to do, in parts, with a paper by Ellsaesser *et al.* (2014), which questions their applicability in practice. In this paper I showed that the critique they put forward is overly generic and not able to target causal graph modeling specifically. Contrary to their arguments, graphs posses a distinctive advantage in revealing when unobserved variables pose a problem for causal effect identification and when assuming conditional independence is justified by theoretical arguments. Exactly these features of transparency and explicitness with respect to the assumptions underlying causal inference determine their value for practical empirical work.

Given the unwarranted criticism of causal graphs by Ellsaesser *et al.*, and their advantages that have been discussed, I hope that this paper will promote further diffusion of causal graph modeling in the field of strategic management. Indeed, directed acyclic graphs have been frequently used by management scholars to visualize and communicate their theories. I encourage them—as other disciplines have done before—to also recognize the powerful empirical methodology DAGs provide. Although theoretical results in the field of graph theory and mathematical statistics can sometimes be difficult to grasp, convenient software tools developed by the computer science and artificial intelligence community allow to benefit from this body of work without the need for extensive formal training. Moreover, these days good introductory-level textbooks exist (Morgan and Winship, 2007; Pearl *et al.*, 2016) that do not require a lot of prior knowledge and offer researchers an easy access to the topic.

## REFERENCES

Alemi F, Zargoush M, Vang J. 2016. Using observed sequence to orient causal networks. *Health Care Management Science*, 1–10.

Angrist JD, Pischke JS. 2009. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Bollen KA, Pearl J. 2013. Eights myths about causality and structural equation models. In Morgan SL (ed.), Handbook of Causal Analysis for Social Research.

Cohen WM, Levinthal DA. 1990. Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, **35**(1): 128–152.

Donaldson L, Qiu J, Luo BN. 2013. For rigour in organizational management theory research. *Journal of Management Studies*, **50**(1): 153–172.

Durand R, Vaara E. 2009. Causation, counterfactuals, and competitive advantage. *Strategic Management Journal*, **30**: 1245–1264.

Ellsaesser F, Tsang EWK, Runde J. 2014. Models of causal inference: impefect but applicable is better than perfect but inapplicable. *Strategic Management Journal*, **35**: 1541–1551.

Elwert F. 2013. Graphical causal models. In Morgan SL (ed.), Handbook of causal analysis for social research, Springer.

Gow ID, Larcker DF, Reiss PC. 2016. Causal inference in accounting research. *Journal of Accounting Research*, **54**(2): 477–523.

Heckman JJ. 2005. The Scientific Model of Causality. *Sociological Methodology*, **35**: 1–97.

Heckman JJ, Pinto R. 2013. Causal Analysis after Haavelmo. *Econometric Theory*, **31**: 115–151.

Imai K, Keele L, Tingley D, Yamamoto T. 2011. Unpacking the Black Box of Causality: Learning about Causal Mechanisms from Experimental and Observational Studies. *The American Political Science Review*, **105**(4): 765–789.

Imai K, Keele L, Yamamoto T. 2010. Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Staistical Science*, **25**(1): 51–71.

Imbens GW. 2004. Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review. *The Review of Economics and Statistics*, **86**(1): 4–29.

Imbens GW, Angrist JD. 1994. Identification and Estimation of Local Average Treatment Effects. *Econometrica*, **62**(2): 467–475.

Kalisch M, Mächler M, Colombo D. 2012. Causal Inference Using Graphical Models with the R Package pcalg. *Journal of Statistical Software*, **47**(11): 1–26.

Kline RB. 2015. Principles and Practice of Structural Equation Modeling. The Guilford Press, New York, NY, 4th ed.

Lewbel A. 2016. The Identification Zoo - Meanings of Identification in Econometrics, working paper.

Loch CH, Wu Y. 2007. Behavioral Operations Management. In Foundations and Trends in Technology, Information and Operations Management, vol. 1, 121–232.

Morgan SL, Winship C. 2007. Counterfactuals and Causal Inference: Methods and Principles for Social Research. Cambridge University Press, 1 ed.

Nelder JA, Wedderburn RWM. 1972. Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)*, **135**(3): 370–384.

Neuberg LG. 2003. Review of 'Causality: Models, Reasoning, and Inference'. *Econometric Theory*, **19**: 675–685.

Pearl J. 1993. Comment: Graphical Models, Causality and Intervention. *Statistical Science*, **8**(3): 266–269.

Pearl J. 1995. Causal diagrams for empirical research. *Biometrika*, **82**(4): 669–709.

Pearl J. 2009. Causality: Models, Reasoning, and Inference. Cambridge University Press, New York, United States, NY, 2nd ed.

Pearl J. 2012. Handbook of Structural Equation Modeling, Guilford Press, New York, NY, chap. The Causal Foundations of Structural Equation Modeling.

Pearl J. 2013. Reflections on Heckman and Pinto's 'Causal analysis after Haavelmo'. Tech. Rep. R-420, Univiversity of California, Los Angeles.

Pearl J. 2014. Interpretation and Identification of Causal Mediation. *Psychological Methods*, **19**(4): 459–481.

Pearl J, Glymour M, Jewell NP. 2016. Causal Inference in Statistics: A Primer. John Wiley & Sons Ltd, West Sussex, United Kingdom.

Powell TC, Lovallo D, Caringal C. 2006. Causal Ambiguity, Management Perception, and Firm Performance. *The Academy of Management Review*, **31**(1): 175–196.

Shpitser I, Pearl J. 2006. Identification of Joint Interventional Distributions in Recursive Semi-Markovian Causal Models. In Twenty-First National Conference on Artificial Intelligence.

Spiegler R. 2016. Bayesian networks and boundedly rational expectations. *The Quarterly Journal of Economics*, **131**(3): 1243–1290.

Spirtes P, Glymour C, Scheines R. 2001. Causation, Prediction, and Search. The MIT Press, Cambride, MA, 2nd ed.

Textor J, Hardt J, Knüppel S. 2011. DAGitty: A Graphical Tool for Analyzing Diagrams. *Epidemiology*, **5**(22): 745.

Vancouver JB, Carlson BW. 2015. All Things in Moderation, Including Tests of Mediation (at Least Some of the Time). *Organizational Research Methods*, **18**(1): 70–91.

VanderWeele TJ, Robins JM. 2007. Four Types of Effect Modification: A Classification Based on Directed Acyclic Graphs. *Epidemiology*, **18**: 561–568.

Varian HR. 2014. Big Data: New Tricks for Econometrics. *Journal of Economic Perspectives*, **28**(2): 3–28.

Weinberg CR. 2007. Can DAGs Clarify Effect Modification? *Epidemiology*, **18**(5): 569–572.

Wright S. 1921. Correlation and Causation. *Journal of Agricultural Research*, **20**: 557–585.

Wright S. 1934. The Method of Path Coefficients. *Annals of Mathematical Statistics*, **5**: 161–215.

Yang J, Zhou Y. 2013. Credit Risk Spillovers Among Financial Institutions Around the Global Credit Crisis: Firm-Level Evidence. *Management Science*, **59**(10): 2343–2359.

Yang Z, Zhou Y. 2016. Quantitative easing and volatility spillovers across countries and asset classes. *Management Science*, **63**(2): 333–354.