Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Operational Research 156 (2004) 92–109
EUROPEAN JOURNAL OF OPERATIONAL RESEARCH
www.elsevier.com/locate/dsw

Production, Manufacturing and Logistics

Causal modeling alternatives in operations research: Overview and application

Ronald D. Anderson [a,1], Gyula Vastag [b,*]

[a] Kelley School of Business, Indiana University, 801 West Michigan Street, BS4053, Indianapolis, IN 46202-5151, USA
[b] Kelley School of Business, Indiana University, 801 West Michigan Street, BS4027, Indianapolis, IN 46202-5151, USA

Received 1 July 2002; accepted 26 November 2002

Abstract

This paper uses the relationships between three basic, fundamental and proven concepts in manufacturing (resource commitment to improvement programs, flexibility to changes in operations, and customer delivery performance) as the empirical context for reviewing and comparing two casual modeling approaches (structural equation modeling and Bayesian networks). Specifically, investments in total quality management (TQM), process analysis, and employee participation programs are considered as resource commitments. The paper begins with the central issue of the requirements for a model of associations to be considered causal. This philosophical issue is addressed in reference to probabilistic causation theory. Then, each method is reviewed in the context of a unified causal modeling framework consistent with probabilistic causation theory and applied to a common dataset. The comparisons include concept representation, distribution and functional assumptions, sample size and model complexity considerations, measurement issues, specification search, model adequacy, theory testing and inference capabilities. The paper concludes with a summary of relative advantages and disadvantages of the methods and highlights the findings relevant to the literature on TQM and on-time deliveries.

Keywords: Manufacturing; TQM; Delivery performance; Causal modeling; Structural equation modeling; Bayesian networks

1. Introduction

Interest in causal modeling methodologies in the social sciences stems from the desire to establish patterns of regularities or laws analogous to those in the physical sciences. A fundamental appeal of causal modeling is the ability to combine cause–effect information, based on theoretical construction, with statistical data to provide a quantitative assessment of relationships among the studied variables. The purposes for employing causal modeling in the study of operations are to develop an explanation of relationships and to provide a basis for inference. The portrayal, evaluation and summarization of assumed causal relationships are the components of explanation. These relationships are then used to develop

* Corresponding author. Tel.: +1-317-278-3681; fax: +1-317-274-3312/278-3312.
E-mail addresses: randers@iupui.edu (R.D. Anderson), gvastag@iupui.edu (G. Vastag).
[1] Tel.: +1-317-274-2446; fax: +1-317-274-3312.

0377-2217/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/S0377-2217(02)00904-9

inferences for diagnostic reasoning from effects to causes and for the prediction of outcomes that would follow from a policy or procedure intervention. Available modeling methods offer differing functional advantages and limitations. However, any method should have potential managerial usefulness by providing outputs with clear interpretation and the capability to assess the impact of potential changes in the modeled process.

Ideally, a causal study would take the form of a randomized controlled experiment conducted over an appropriate time period. Such a research design would minimize construct, internal, external, and statistical threats to validity (Cook and Campbell, 1979), and allow the possibility of causal conclusions to be reached. Unfortunately, randomized controlled experiments can seldom, if ever, be utilized to provide causal knowledge for strategy and policy issues. Thus, causal modeling methods for non-experimental data are of interest.

Bayesian networks and structural equation models (SEM) are the causal modeling methods for non-experimental data reviewed and compared in this paper. The paper begins with the central issue of the requirements for a model of associations to be considered causal. This philosophical issue is addressed in reference to probabilistic causation theory. Then, each method is reviewed in the context of a unified causal modeling framework consistent with probabilistic causation theory, and applied to a common dataset. The comparisons include concept representation, distribution and functional assumptions, sample size and model complexity, measurement, specification search, model adequacy, theory testing and inference capabilities. The paper concludes with a summary of the relative advantages and disadvantages of the methods.

2. Probabilistic causation theory

The area of causation has been extremely active over the past twenty years with numerous interactions between the fields of philosophy, statistics and computer science. This activity has spawn spirited controversy on a wide variety of conceptual and methodological issues (McKim and Turner, 1997). Causality, as a theoretical postulate, has been the subject of highly contested discussions since the reductive account offered by Hume (1969). Hume characterized causation by the regularity of constantly conjoined pairs of events (Effect $= f(\text{Cause})$), under conditions of temporal priority (a cause must precede an effect), and contiguity (a cause is temporally adjacent to an effect). However, Hume's account does not provide for imperfect regularities nor does it have the ability to distinguish between a genuine causal relation and a spurious association. These weaknesses motivated development of theories of causation that cast causal relationships between general events in terms of stochastic descriptions (Suppes, 1970).

The key feature of probabilistic causation is a paradigm switch from the absolute determination of an effect due to the occurrence of a cause to the occurrence of a cause increasing the probability of an effect. An assumption underlying this perspective is that incomplete knowledge of causes results in uncertain cause–effect relationships. This conceptualization, labeled as pseudo-indeterminism (Spirtes et al., 1993), assumes that specified causes do not alone determine an effect, but do so in conjunction with unspecified unobserved causes. Thus, pseudo-indeterminism assumes that sets of independent specified causes and unspecified causes are the direct causes ($\rightarrow$) of an effect: specified causes $\rightarrow$ effect $\leftarrow$ unspecified causes.

Cause–effect relationships, under the assumption of pseudo-indeterminism, may be encoded into a graphical structure known as a directed acyclic graph or simply a DAG. Each arrow in a DAG depicts causal dependence and the absence of a connecting arrow indicates causal independence. The encoded structure is characterized as directed, since two-headed arrows depicting non-causal association are not allowed and as acyclic, since feedback loops (e.g., $X \rightarrow Y \rightarrow X$) are not allowed.

The common cause principle states if two variables in a population are associated and neither is a cause of the other, they must share a common cause (Reichenbach, 1956). The term association is used, in reference to probabilistic dependence