

Chapter 5

CAUSALITY AND STRUCTURAL MODELS IN SOCIAL SCIENCE AND ECONOMICS

*Do two men travel together
unless they have agreed?*

Amos 3:3

Preface

Structural equation modeling (SEM) has dominated causal analysis in economics and the social sciences since the 1950s, yet the prevailing interpretation of SEM differs substantially from the one intended by its originators and also from the one expounded in this book. Instead of carriers of substantive causal information, structural equations are often interpreted as carriers of probabilistic information; economists view them as convenient representations of density functions, and social scientists see them as summaries of covariance matrices. The result has been that many SEM researchers are having difficulty articulating the causal content of SEM, and the most distinctive capabilities of SEM are currently ill understood and under utilized.

This chapter is written with the ambitious goal of reinstating the causal interpretation of SEM. We shall demonstrate how developments in the areas of graphical models and the logic of intervention can alleviate the current difficulties and thus revitalize structural equations as the primary language of causal modeling. Toward this end, we recast several of the results of Chapters 3 and 4 in parametric form (the form most familiar to SEM researchers) and demonstrate how practical and conceptual issues of model testing and parameter identification can be illuminated through graphical methods. We then move back to non-parametric analysis, from which an operational semantics will evolve that offers a coherent interpretation of what structural equations are all about (Section 5.4). In particular, we will provide answers to the following fundamental questions: What do structural equations claim about the world? What portion of those claims is testable? Under what conditions can we estimate structural parameters through regression analysis?

In Section 5.1 we survey the history of SEM and suggest an explanation for the current erosion of its causal interpretation. The testable implications of structural models are explicated in Section 5.2. For recursive models (herein termed *Markovian*), we find that the statistical content of a structural model can be fully characterized by a set of zero partial correlations that are entailed by the model. These zero partial correlations can be read off the graph using the *d-separation* criterion, which in linear models, applies to graphs with cycles and correlated errors as well (Section 5.2). The application of this criterion to model testing is discussed in Section 5.2.2, which advocates local over global testing strategies. Section 5.2.3 provides simple graphical tests of model equivalence and thus clarifies the *nontestable* part of structural models.

In Section 5.3 we deal with the issue of determining the identifiability of structural parameters prior to gathering any data. In Section 5.3.1, simple graphical tests of identifiability are developed for linear Markovian and semi-Markovian models (i.e., acyclic diagrams with correlated errors). These tests result in a simple procedure for determining when a path coefficient can be equated to a regression coefficient and, more generally, when structural parameters can be estimated through regression analysis. Section 5.3.2 discusses the connection between parameter identification in linear models and causal effect identification

in nonparametric models, and Section 5.3.3 offers the latter as a semantical basis for the former.

Finally, in Section 5.4 we discuss the logical foundations of SEM and resolve a number of difficulties that were kept dormant in the past. These include operational definitions for structural equations, structural parameters, error terms, and total and direct effects, as well as a causal-theoretic explication of exogeneity in econometrics.