Sept., 1981

COMPLEXITIES OF MEDIATION STRUCTURE; or, A FUNNY THING HAPPENED ON THE WAY TO MODA'S REVISION

Wm. W. Rozeboom

The attached document began life as the second draft of the monograph on applied data analysis previewed in the first few pages here. But it has evolved into an in-depth study of certain limited but central aspects of causal structure as this has traditionally been viewed in the multivariate literature. Because revision of Chapter 1 was largely completed before Chapter 2 took its present explosive turn, some practical preliminaries therein are no longer relevant just now.

The funny thing that happened was my jolting discovery that a fundamental principle of mediated causality which had long seemed obvious and indisputable to me is in fact not generally true. (See Fallacious Thesis 1, p. 2.22.) I know of no published work that explicitly promulgates this fallacy, but that is only because scarcely any explications of our behavioral-science intuitions about causality have ever appeared. Nevertheless, I submit that it is implicit throughout modern multivariate analysis, especially in recursive causal modelling, which is where questions of causal composability are most salient. Although the additional model premises needed to justify our standard intuitions in this respect usually seem plausible enough in particular applications, the deeper point is that significant additional premises are in fact needed. Surely it is thus reasonable to contend that the current burgeoning of causal modelling in applied data analysis--and why else bother to parse data if not to seek information about causal principles that underlie them?--creates an increasingly urgent need for a theory of mediation structure that gives us some notion of what we are talking about when we propose causal models and estimate parameters in conjectured structural equations. A foundation for that limited theory (i.e., for just the logic of causal mediation/composition, not for other facets of causal structure also much in need of exposition) is herewith proposed.

Unhappily, the account that emerges here is densely technical. Qualitatively, it can be viewed as no more than a rationalized extension of orthodox causal-path modelling. But to build that rationalization, and to develop from it an analysis of

the conditions under which the composition of one causal regularity into another is itself a (mediated) causal regularity, has required creation of a new language, or rather, what is even more demanding, a new mathematical system. Chapter 2 draws out the exact logical force of certain basic causal-structure postulates that have ample intuitive justification, but which encorporate axiomatic concepts that are not usually understood in the precisely detailed senses given to them here. And additional terms are introduced by explicit definitions that make good sense locally but globally become burdensome to memory. Moreover, apart from the digraph representation of direct-source relations and some elementary algebraic notions, the reader will not likely have experienced mathematical constructions sufficiently similar to the present ones to effect much positive transfer of comprehension--one has to work up here pretty much from scratch.

In short, even were this optimally written, the nature of the material would make it tough slogging--not because it is inherently difficult, but because it demands patient attention to the details of novel concepts. Moreover, an extra barrier to comprehension of this initial draft is just that it <u>is</u> an initial draft. Undoubtedly there are many needless obscurities in the present exposition (especially in its proofs, which I have tried to keep short but may well have over-compressed) that with effort can be cleaned out. Some improvements I can manage on my own; but to do the job right I need constructive feedback about what in this is most opaque/confusing/wrong.

On the other hand, regardless of improvements, perhaps the material here is just too recondite, given the present state-of-the-art in multivariate thinking, to be publishable in anything like its present form. Even if you find the details of this incomprehensible, are you able to acquire from overview skimming some feeling for the problems it addresses, the general direction of its development, and the shape of its achievements? If so, have you any advice for me on the feasibility of seeking to publish fragments of this material? Possibly it is best written off as hopeless.

September, 1981

MULTIPLE-OUTPUT DEPENDENCY ANALYSIS: THE STEP BEYOND MANOVA

William W. Rozeboom

University of Alberta

CHAPTER 1. CLASSICAL MODELS OF CAUSALITY.

Introduction.

Scientific data can arise in many diverse formal patterns, the generic study of which is a chapter in the logic of data analysis still to be written. Yet one particular form has so dominated research practice that data structured this way may be thought of as "classical." Specifically, let us say that a classical data array is an $n \times N$ numerical score matrix $\underset{\sim}{Z}$ whose rows correspond to an n-tuple $\underset{\wedge}{Z} = \langle \underset{\wedge}{z}_1, \ldots, \underset{\wedge}{z}_n \rangle$ of numerically scaled data variables, whose columns correspond to the N members of a sequentially indexed sample population P of subjects (i.e., P is an N-tuple of observational/experimental units), and whose ijth element z_{ij} is the score on variable $\underset{\wedge}{z}_i$ observed for the jth subject in P. Insomuch as initial data records can generally be parsed in more than one way, classical data arrays result in part from our decision to express our raw observations in this particular format. Obviously there are a number of formally different but transformationally equivalent structures that could replace this definiens (notably, transposing Z's rows and columns, as preferred by some data analysts), all of which may be viewed as classical in an abstract sense. But for present purposes, numerical Variables-by-Subjects score matrices are the most convenient embodiment of the abstract classical data form.

Since any variable can be scaled numerically, no matter how qualitative the attributes represented by its alternative scale values, classical data arrays can include variables of any substantive character. In particular, these can be metrical (quantitative), categorical (qualitative), or a mix of both. And incomplete Variables-by-Subjects data arrays can often be treated as classical by filling the

empty cells with artificial scores construed as estimates of missing data. (See Rozeboom, [GLDA], for details on this point.) Even so, not all data structures that arise in modern research are classical in the sense just defined. Repeated-measurements data, wherein individuals are observed on the same measures across two or more occasions, have a structure more complex than the classical form; and so do data on polyadic variables whose values represent relations among two or more subjects. (In contrast, classical data typically involve just monadic variables whose values represent subjects' nonrelational properties.) Although repeated-measurements and polyadic data can be transformed into classical arrays--quite appropriately so for some aspects of their analysis--the theory of their interpretation goes beyond their reduction to classical form. (Rozeboom, [GLDA], discusses multivariate repeated measurements in operational detail; while the nature of polyadic data analysis is briefly sketched in Rozeboom, 1966 pp. 198-214, 1972 pp. 111-114.) Even so, not only do classical data arrays still prevail in the behavioral sciences, the logic of their analysis generally underlies that of more complex data structures. So while present concerns explicitly address only the classical case, that does not strongly limit the applicability of whatever new operational techniques or deepened theoretical insights may emerge here.

What to do with classically structured data.

Many algorithms have been devised to extract information from classical data arrays, but the ones most deserving of interpretive respect are those that under favorable circumstances identify parameters of the data's causal production. Extant varieties of causal analysis divide between two historically disjoint approaches: (1) In traditional dependency analysis, notably Analysis of Variance/Covariance (ANOVA/ANCOVA) and various regression models, one or more of the study's data variables are considered to be causally affected by some or all of the others; and under presumption that these dependencies are of a generic algebraic form Φ, the analysis extracts from the data matrix an estimate of the specific Φ-form function by which

each dependent data variable is determined in the sample population by its putative observed sources. In this tradition's most recent extension, Multivariate Analysis of Variance (MANOVA), a multiplicity of dependent data variables is treated as a space that the analyst prefers to span by orthogonal axes thereof that are successively most predictable from the independent data variables; but otherwise, MANOVA's aim is no different from that of ANOVA/ANCOVA. (2) In contrast, the tradition of inter-dependency analysis (a useful contrastive label even though technically a misnomer) treats all the data in a classical array as symptoms of shared underlying causes, and the analysis seeks to learn whatever it can about these unobserved source factors from the data variables' joint sample distribution. The most broadly powerful version of this approach is linear inferential factor analysis; but other, more limited, models for hidden-source recovery have also appeared, notably Latent Structure Analysis (cf. Lazarsfeld & Henry, 1968) for dealing with multivariate categorical distributions.

As these two styles of causal analysis have advanced to date, each has needlessly retained an important early deficiency in what happens to be the other's special strength. On the side of dependency analysis, which is the core of psychology's experimental tradition (cf. Cronbach, 1959), it is still uncommon to study more than one or two dependent variables at a time; yet if one's research design does include a decent multiplicity of dependent variables, it is unconscionably wasteful of the data's information content to ignore the implications of whatever output covariation therein is unexplained by the observed independent variables. And across the aisle, it is equally parochial for correlational psychology (again cf. Cronbach, 1959) not to explore whether the multidimensional symptom configurations at issue in inter-dependency studies might be due in part to common sources that can be externally manipulated or at least observed directly.

In fact, it turns out that apart from some softening of the statistical tests that enthrall ANOVA partisans, the standard logic of inferential factor analysis can be combined almost effortlessly with dependency analysis's generic linear model for estimating the force of observed causes. Scarcely any new techniques are required;

one needs only to perceive how the old methodologies mesh. The theory of this union is herewith developed under the acronym MODA, for Multiple-Output Dependency Analysis.

MODA operates upon classical data arrays in which the observed variables are partitioned into three disjoint groups: (1) output (dependent) variables, (2) input (independent) variables presumed to be causally antecedent to these outputs, and (3) indicators ("covariates") thought to be diagnostic of additional source factors for which we would like to control when inferring input/output dependencies governing these data. Any number of output variables greater than zero is acceptable, the more the better, so long as these are metrical or, with caution, binaries. (Categorical outputs are generally unsuitable for MODA, but binary outputs can be tolerated so long as they are all logically independent of one another.) In contrast, the input and indicator groups can be empty, and it is best that neither be overabundant. MODA's only constraints on the character of inputs and indicators are that (i) any categorical dimensions therein must be recast as a tuple of binaries such as correspond to "dummy variable" codings of ANOVA main-effect and interaction components, and (ii) the combined input/indicator sample covariance matrix must not have any near-zero roots. When there is only one output variable, MODA reduces to orthodox regression analysis or, if the inputs include binary codings of treatment categories, to ANOVA/ANCOVA. And at the other boundary of its range, when inputs and indicators are both null, MODA becomes ordinary factor analysis.

Computationally, MODA proposes little that is at all unfamiliar--which is why scarcely anything will be said here about computational detail. Rather, concern is for the rationale of analyzing classical data arrays one way rather than another when options exist, with special attention to background issues seldom aired in the multivariate literature. Present treatment of these matters is often brusk, fragmentary, and undoubtedly controversial; but at least the issues are raised. If you don't like what I say, or are distressed by omissions, then by all means let's talk further. Promotion of such discussion is a major objective of this essay.

Notation and special terminology.

A numerically scaled scientific variable -- henceforth simply "variable" -- over a population $\underline{P}$ is a function mapping each member of $\underline{P}$ into a number that represents some property of that subject. (For details, see Rozeboom, 1966b, pp. 175-181.) Finite ordered sets, i.e. "tuples," of variables over a common population $\underline{P}$ are denoted here by capital script letters $\underset{\wedge}{X}$, $\underset{\wedge}{Y}$, $\underset{\wedge}{Z}$, $\underset{\wedge}{E}$, $\underset{\wedge}{F}$, $\underset{\wedge}{G}$, and $\underset{\wedge}{H}$, the subscript-indexed lowercase counterparts of which stand for single variables in those tuples. Thus, $\underset{\wedge}{X}$ is a tuple of variables $\langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_m \rangle$, $\underset{\wedge}{Y}$ a tuple $\langle \underset{\wedge}{y}_1, \ldots, \underset{\wedge}{y}_n \rangle$, etc. Defining $\underset{\wedge}{X}$ (etc.) to be the ordered set of variables $\langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_m \rangle$ leaves open how that order is to be expressed vectorially when $\underset{\wedge}{X}$ enters matrix equations. Here, tuples of variables are always taken algebraically to be column vectors. Correspondingly, we express scores on variables $\underset{\wedge}{X}$ (etc.) for members of $\underline{P}$ by a Variables-by-Subjects matrix $\underset{\wedge}{X}$ whose $\underline{ij}$th element is the score of subject $\underline{j}$ on the $\underline{i}$th variable in tuple $\underset{\wedge}{X}$. In all equations below, tuples of variables can be replaced by the corresponding score matrices in $\underline{P}$ so long as the replacement is uniform and additive constants are appropriately expanded into matrices. Whenever possible, we impose the constraint that any tuple $\underset{\wedge}{X}$ of variables contains no repeated elements, i.e., that $\underset{\wedge}{x}_i \neq \underset{\wedge}{x}_j$ if $\underline{i} \neq \underline{j}$. Accordingly, when $\underset{\wedge}{X} = \langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_m \rangle$ and $\underset{\wedge}{Y} = \langle \underset{\wedge}{y}_1, \ldots, \underset{\wedge}{y}_n \rangle$, we shall understand $\langle \underset{\wedge}{X}, \underset{\wedge}{Y} \rangle$ to be the subtuple of $\langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_m, \underset{\wedge}{y}_1, \ldots, \underset{\wedge}{y}_n \rangle$ comprising just the distinct variables therein in the order of their first occurrence. Notation $\underset{\wedge}{X} = \langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_n \rangle$ for the variables in tuple $\underset{\wedge}{X}$ implies that there are $\underline{m}$ distinct variables in $\underset{\wedge}{X}$ and hence that for all $\underline{i}, \underline{j} = 1, \ldots, \underline{m}$, $\underset{\wedge}{x}_i \neq \underset{\wedge}{x}_j$ if $\underline{i} \neq \underline{j}$. Note that "subtuple" is more general than "proper subtuple, i.e., one subtuple of $\underset{\wedge}{X}$ is $\underset{\wedge}{X}$ itself.

With one exception, matrices of covariances between tuples of variables will be written as sans serif capital $\underset{\sim}{C}$ with a double subscript indicating which variables correspond to rows and which to columns. Thus, $\underset{\sim}{C}_{YX}$ is the matrix whose $\underline{ij}$th element is the covariance $c_{y_i x_j}$ between the $\underline{i}$th variable in $\underset{\wedge}{Y}$ and the $\underline{j}$th variable in $\underset{\wedge}{X}$. For covariance matrices stipulated to be diagonal, $\underset{\sim}{C}$ will be replaced by $\underset{\sim}{D}$. Sans serif capital $\underset{\sim}{D}$, subscripted as appropriate, will always denote a matrix that is

diagonal. More generally, matrices of real numbers (the only matrices that occur here) will be expressed by sans serif capital letters, real vectors usually by lowercase sans serif letters, and individual real numbers usually by lowercase italic letters. But also, for any matrix $\underset{\sim}{A}$, $[\underset{\sim}{A}]_{ij}$ is the ijth element of $\underset{\sim}{A}$, while $[\underset{\sim}{A}]_{i.}$ and $[\underset{\sim}{A}]_{.j}$ are respectively the ith row and the jth column of $\underset{\sim}{A}$. Similarly, $[\underset{\sim}{a}]_i$ is the ith element of vector $\underset{\sim}{a}$. And we will sometimes write $\underset{\sim}{Z}$ (etc.) for an arbitrary tuple of scores on variables $\underset{\sim}{Z}$. If matrices $\underset{\sim}{A}$ and $\underset{\sim}{B}$ are respectively $m \times r$ and $m \times s$, $[\underset{\sim}{A} \ \underset{\sim}{B}]$ is the $m \times (r+s)$ partitioned matrix comprising $\underset{\sim}{A}$ as its first r

columns and B as its last $\underline{s}$ columns. Extension of this notation to other matrix partitionings will be obvious.

Unlike prevailing custom in the modern multivariate literature, basic vectors (i.e., ones not marked transpose) are here allowed to be either row vectors or column vectors as specified by context, with row vectors the more common. This departure from notational orthodoxy is dictated by two desiderata: (1) to write matrix equations in standard algebraic format wherein coefficients premultiply the variables to which they apply, and (2) not to proliferate transposition operations needlessly nor to fight against the flow of notation that follows most naturally from satisfaction of (1). To illustrate, suppose that each variable in tuple $\underset{\wedge}{Y}$ is a linear function of variables $\underset{\wedge}{X}$. To express this relation in compliance with desideratum (1) we write

$$\underset{\wedge}{Y} = \underset{m}{a} + \underset{mYX}{B}\underset{\wedge}{X} ,$$

where X and Y are column vectors of variables $\langle x_1,\ldots,x_m\rangle$ and $\langle y_1,\ldots,y_n\rangle$, respectively, $\underset{m}{a}$ is an order-$\underline{m}$ column vector of additive constants, and $\underset{mYX}{B}$ is an $\underline{m}\times\underline{n}$ coefficient matrix to which the double subscript is appended to point out that the rows and columns of $\underset{mYX}{B}$ correspond to the variables in $\underset{\wedge}{Y}$ and in $\underset{\wedge}{X}$, respectively. It would be perverse to write this equation as $\underset{\wedge}{Y} = \underset{m}{a} + \underset{mXY}{B'}\underset{\wedge}{X}$. But it is similarly perverse to write $\underset{\wedge}{y} = \underline{a} + \underset{mXy}{b'}\underset{\wedge}{X}$ for its special case where $\underset{\wedge}{Y}$ comprises just one variable $\underset{1}{y}$. Instead, natural notation for the latter is

$$\underset{1}{y} = \underline{a} + \underset{myX}{b}\underset{\wedge}{X} ,$$

even though that requires the basic coefficient vector to be a row. Insistence that all basic vectors appear in matrix equations as columns incurs sufficient loss of other algebraic conveniences and familiarities that it seems worthwhile to test whether relaxation of local custom in this regard may not improve upon the efficiency or at least comfort of the currently prevailing format.

This essay presumes that its readers are modestly familiar with covariance, multiple regression, inferential factor analysis, and, if occasional asides are to

be understood, ANOVA/ANCOVA. Even so, it is prudent to review the elementary properties of covariances that will be repeatedly exploited later. Let $\underset{\sim}{J}_N$ be defined as the $\underline{N}$x $\underline{N}$ row-centering matrix

$$\underset{\sim}{J}_N =_{def} \underset{\sim}{I}_N - \underline{N}^{-1}\underset{\sim}{1}'_N\underset{\sim}{1}_N$$

in which $\underset{\sim}{1}_N$ is the order-$\underline{N}$ Unity row vector, i.e. $\underset{\sim}{1}_N = \langle 1,1,\ldots,1,1\rangle$, and $\underset{\sim}{I}_N$ is the $\underline{N}$x $\underline{N}$ Identity matrix. Order-subscript $\underline{N}$ will henceforth be omitted from $\underset{\sim}{1}_N$, $\underset{\sim}{I}_N$, and $\underset{\sim}{J}_N$ whenever the particular $\underline{N}$ involved is either irrelevant or implied by context. Observe that $\underset{\sim}{J}$ is symmetric, idempotent, and annihilates $\underset{\sim}{1}$, i.e.,

$$J' = J, \quad \underset{\sim}{J}^2 = \underset{\sim}{J}, \quad \underset{\sim}{1}\underset{\sim}{J} = \underset{\sim}{0},$$

and that for the $\underline{N}$-tuple $\underset{\sim}{z}$ of scores in population $\underline{P}$ on any variable $\underset{\wedge}{z}$, since the mean of $\underset{\wedge}{z}$ in $\underline{P}$ is

$$\underline{m}_z =_{def} \underline{N}^{-1}\underset{\sim}{z}\underset{\sim}{1}',$$

postmultiplying $\underset{\sim}{z}$ by $\underset{\sim}{J}$ (i.e. $\underset{\sim}{z}\underset{\sim}{J} = \underset{\sim}{z} - \underline{m}_z\underset{\sim}{1}$) deviates each score in $\underset{\sim}{z}$ from $\underset{\wedge}{z}$'s mean in $\underline{P}$. Hence the covariance between any variables $\underset{\wedge}{y}$ and $\underset{\wedge}{x}$ in $\underline{P}$ is

$$\underline{c}_{yx} =_{def} \underline{N}^{-1}(\underset{\sim}{y}\underset{\sim}{J})(\underset{\sim}{x}\underset{\sim}{J})' = \underline{N}^{-1}\underset{\sim}{y}\underset{\sim}{J}\underset{\sim}{J}'\underset{\sim}{x}' = \underline{N}^{-1}\underset{\sim}{y}\underset{\sim}{J}\underset{\sim}{x}',$$

while more generally, the matrix of covariances in $\underline{P}$ between the variables in any two tuples $\underset{\wedge}{Y}$ and $\underset{\wedge}{X}$ is

$$\underset{\sim}{C}_{YX} = \underline{N}^{-1}(\underset{\sim}{Y}\underset{\sim}{J})(\underset{\sim}{X}\underset{\sim}{J})' = \underline{N}^{-1}\underset{\sim}{Y}\underset{\sim}{J}\underset{\sim}{X}'.$$

It follows immediately that if variables $\underset{\wedge}{Y}$ and $\underset{\wedge}{X}$ are linear functions in $\underline{P}$ of factor tuples $\underset{\wedge}{F}$ and $\underset{\wedge}{G}$, respectively, i.e., if

$$\underset{\wedge}{Y} = \underset{\sim}{a}_Y + \underset{\sim}{B}_{YF}\underset{\wedge}{F}, \quad \underset{\wedge}{X} = \underset{\sim}{a}_X + \underset{\sim}{B}_{XG}\underset{\wedge}{G},$$

then since $\underset{\sim}{Y}\underset{\sim}{J} = (\underset{\sim}{a}_Y\underset{\sim}{1} + \underset{\sim}{B}_{YF}\underset{\sim}{F})\underset{\sim}{J} = \underset{\sim}{B}_{YF}(\underset{\sim}{F}\underset{\sim}{J})$ and similarly for $\underset{\wedge}{X}$, we have

$$\underset{\sim}{C}_{YY} = \underset{\sim}{B}_{YF}\underset{\sim}{C}_{FF}\underset{\sim}{B}'_{YF}, \quad \underset{\sim}{C}_{XX} = \underset{\sim}{B}_{XG}\underset{\sim}{C}_{GG}\underset{\sim}{B}'_{XG}, \quad \underset{\sim}{C}_{YX} = \underset{\sim}{B}_{YF}\underset{\sim}{C}_{FG}\underset{\sim}{B}'_{XG}.$$

And obviously, for any two tuples $\underset{\wedge}{Y}$ and $\underset{\wedge}{X}$, $\underset{\sim}{C}'_{YX} = \underset{\sim}{C}_{XY}$.

The logic of same-subject causality.

In classical dependency analysis, one observes for each member of sample population P a score on some output (causally dependent) variable $\underset{\wedge}{y}$ together with scores on one or more independent variables $\underset{\wedge}{X} = \langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_n \rangle$ that are thought to be sources (causal determinants) of $\underset{\wedge}{y}$. The data analyst's task is then to diagnose from this score array (1) the lawful regularity (causal function) by which $\underset{\wedge}{X}$ contributes to y's determination in the population P* sampled by P; (2) the degree of predictability with which $\underset{\wedge}{y}$ is determined in P* just by $\underset{\wedge}{X}$; and (3) how reliable are the estimates of (1) and (2) we obtain from these data. There is much to be said theoretically about each of these; but many aspects of that theory will be slighted here, (3) in particular. We are concerned mainly with the practicalities of (1), and for the most part shall not find it useful to distinguish the population P for which we have real or hypothetical data from some more inclusive population P* of which P is just a sample. Any method of applied data analysis needs first of all to think through what it would do if it had omniscient access to all particular events from which it aspires to infer governing regularities before it can meaningfully worry how best to proceed from just a fragment of such ideally complete data. Deriving suitable humility about the generalizability of one's results from a proper regard for their sampling uncertainty is indeed an important facet of applied data analysis, and eventually (p. ff. below) we shall face up to MODA's responsibilities in that regard. But our main statistical tradition of null-hypothesis testing is outrageously ill-suited for that purpose; and even the most cogent appraisals based on likelihood functions are distressingly conditional on unargued presuppositions justified more by mathematical convenience than by real-world plausibility (see p. , below). As it is, we shall see that parameters reclaimed by MODA are corrupted by "sampling error" only through within-sample departures from zero of certain covariances that we have no real assurance would vanish even were our sample population arbitrarily large.

Causal hypotheses.

When we envision that variable $\underset{\wedge}{y}$ has causal sources in population P, just what are we conjecturing? This is so deeply complex a matter that even to itemize the more salient issues is impractical here, much less to discuss them adequately. On this occasion I shall attempt to work within the idealized framework that is already familiar in the multivariate literature, with only grudging hints at the mysteries beyond. See Rozeboom ([GLDA]) for a more articulate though still superficial sketch of causal regularity and Rozeboom (in preparation) for my best effort to explore this comprehensively.

As an idealized first-approximation, let us say that a variable $\underset{\wedge}{y}$ is "caused," jointly and deterministically, by variables $\underset{\wedge}{X} = \langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_m \rangle$ in any real population P of subjects just in case (1) there is a function ϕ such that $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ in P, i.e., $y_s = \phi(\underline{X}_s)$ for the scores y_s and $\underline{X}_s$ on $\underset{\wedge}{y}$ and $\underset{\wedge}{X}$, respectively, for each subject s in P, and (2) there exist principles of natural necessitation (de re necessity, not de dicto) and an array $\propto$ of attributes ("boundary conditions") common to the members of P such that for any subject s, s's having attributes $\propto$ and values $\underline{X} = \langle \underline{x}_1, \ldots, \underline{x}_m \rangle$ of variables $\underset{\wedge}{X}$ jointly necessitate that s have value $\phi(\underline{X})$ of variable $\underset{\wedge}{y}$. We shall call any such generality $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$, characterizing the joint causal force of variables $\underset{\wedge}{X}$ for variable $\underset{\wedge}{y}$ under some boundary conditions $\propto$, a "causal regularity" in any population P of subjects satisfying $\propto$. Causal regularity $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ in P is "nomically irreducible" just in case there is no function ϕ' and proper subtuple $\underset{\wedge}{X}'$ of $\underset{\wedge}{X}$ such that $\underset{\wedge}{y} = \phi'(\underset{\wedge}{X}')$ is also a causal regularity in P, and is "functionally irreducible" just in case each variable $\underset{\wedge}{x}_i$ in $\underset{\wedge}{X}$ actually matters for $\phi(\underset{\wedge}{X})$, i.e., just in case for each $\underset{\wedge}{x}_i$ in $\underset{\wedge}{X}$ there exist tuples $\underline{X}$ and $\underline{X}''$ of logically possible joint values of $\underset{\wedge}{X}$ such that $\underline{X}''$ differs from $\underline{X}$ just in the value of $\underset{\wedge}{x}_i$ and $\phi(\underline{X}) \neq \phi(\underline{X}'')$. Strictly speaking, no regularity counts as joint causality unless it is nomically irreducible, since otherwise at least one of its independent variables does not truly conjoin the others in determining $\underset{\wedge}{y}$; so we henceforth presume that no lawlike regularity is strictly causal unless

(See p. 21, below, for additional clarification of this notion.)

it is nomically irreducible. Prima facie the same should be true of functional irreducibility. However, this appearance proves to be deceptive (cf. p. 2.34f. below): and in any case it is often technically advantageous to relax irreducibility requirements. So we shall later speak of "structural equations" in a liberal sense under which a dependency $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ counts as "structural" even when some variables in $\underset{\wedge}{X}$ are irrelevant to the value of $\phi(\underset{\wedge}{X})$, so long as $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ can be reduced to a causal regularity by eliminating variables from $\underset{\wedge}{X}$ that do not work jointly with the rest of $\underset{\wedge}{X}$ in bringing about $\underset{\wedge}{y}$.

Precisely what might be meant by Clause 2 of this definition of "causal regularity" is an obscurity that centuries of philosophers' efforts have done little to clarify. Nevertheless, the notion is indispensable for the conduct of our real-world affairs: It grounds our practical reasoning of form "Our bringing it about that ___ will have consequences ___," and more generally is deeply foundational for the inductive inferences we find intuitively rational (see Rozeboom, 1971, 1981), As empirical scientists and engineers, our responsibility for explicating causality concepts is to make as articulately explicit as we can the force of these notions in our professional thinking, tightening and correcting our intuitions as overt awareness of them reveals confusions and infelicities. Given a sufficiently rich data base in our <u>praxis</u> of causality, philosophic illumination of its nature will surely not be far behind.

The simplified model of causality to be exploited here includes certain assumptions that had best be noted at the outset.

First of all, we treat all causal laws as fully deterministic with a finite number of independent variables. As elaborated later (p.16f.), this is at worst a convenient heurism having no discernably harmful side effects, and for all we know may even be true. But we also recognize that one event, or one tuple of events, can be <u>a</u> cause of another without being a <u>complete</u> cause of it. So when variables $\underset{\wedge}{X}$ are surmised to be joint sources of variable y in population <u>P</u>, our conjecture is in general only that $\underset{\wedge}{y}$ is determined in <u>P</u> jointly by $\underset{\wedge}{X}$ together with some

possibly-null supplementary tuple $\underset{\wedge}{Z}$ of $\underset{\wedge}{y}$-sources. To clarify our notions in this regard, let us say

Definition 1.1. Variables $\underset{\wedge}{X} = \langle x_1,\ldots,x_m\rangle$ are a strictly joint (causal) source of variable $\underset{\wedge}{y}$ in population $\underline{P}$ iff $\underset{\wedge}{y}$ is determined in $\underline{P}$ under some strict (i.e. nomically irreducible) causal regularity $\underset{\wedge}{y} = \phi(\underset{\wedge}{X},\underset{\wedge}{Z})$ whose input variables $\langle\underset{\wedge}{X},\underset{\wedge}{Z}\rangle$ include all of $\underset{\wedge}{X}$. If moreover supplementary tuple $\underset{\wedge}{Z}$ can be taken null, i.e., if $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ is a strict causal regularity in $\underline{P}$ for some ϕ, $\underset{\wedge}{X}$ is (or variables $\underset{\wedge}{X}$ are) a strictly complete (causal) source of $\underset{\wedge}{y}$ in $\underline{P}$. If singleton tuple $\underset{\wedge}{X} = \langle\underset{\wedge}{x}\rangle$ is a strictly joint source of $\underset{\wedge}{y}$ in $\underline{P}$, we say simply that $\underset{\wedge}{x}$ is a (causal) source of $\underset{\wedge}{y}$ in $\underline{P}$.

These varieties of causal-source relations will be expanded in Chapter 2. In particular, we shall there regard a tuple $\underset{\wedge}{X}$ of variables as an "inclusively" complete source of variable $\underset{\wedge}{y}$ in $\underline{P}$ iff some subtuple of $\underset{\wedge}{X}$ is a strictly complete source of $\underset{\wedge}{y}$ in $\underline{P}$.

Secondly, we shall attend only to causal regularities under which events coupled as cause and effect are formalized as having the same subject. That is, for any causal regularity $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ defined as above, the (compound) event of subject $\underline{s}$ in $\underline{P}$ having tuple $\underline{X}$ of values on independent variables $\underset{\wedge}{X}$ causes that same $\underline{s}$ to have value $\phi(\underline{X})$ of dependent variable $\underset{\wedge}{y}$. Although technical data analysis has not yet sought to relax that constraint, it seems absurdly narrow at first blush, insomuch as the spatio-temporal location of a dependent event appears almost always to differ somewhat from the locations of its causes. Thus when John scratches his nose, his itch is not strictly synchronous with the scratching it elicits but slightly prior to it; and the itch in turn may well have been caused in part by properties of objects external to John. Similarly, the state of baby Jane's chromosomes today has no effect on her height right now, but will much influence Jane's height 15 years hence through a chain of causal mediations, and is in turn due largely to the chromatin character of Jane's parents at the time they produced the gametes whose fusion was Jane's origin. Just the same, cause/effect sequences between subjects can generally be subsumed under same-subject laws by various tricks of definition. In the important special case of causation across time differences--"lags"--within the same temporally extended individual (e.g. John throughout his lifetime), for any lag Δ and variable $\underset{\wedge}{x}$ whose subjects are time-slices $\{\underline{i}\text{-at-time-}\underline{t}\}$ of enduring individuals $\{\underline{i}\}$ we can define $\underset{\wedge}{x}^{\Delta}$ to be the variable over population $\{\underline{i}\text{-at-}\underline{t}\}$ whose value for each individual-$\underline{i}$-at-time-$\underline{t}$ is in fact $\underline{i}$'s value of $\underset{\wedge}{x}$ at time $\underline{t} - \Delta$. Then the causal influence of $\underline{i}$'s $\underset{\wedge}{x}$-input at time $\underline{t}-\Delta$ upon $\underline{i}$'s $\underset{\wedge}{y}$-output at time $\underline{t}$ can be taken to instantiate a same-subject dependency of $\underset{\wedge}{y}$ upon $\underset{\wedge}{x}^{\Delta}$. (See Rozeboom, [GLDA], for expansion of this point.) To a large extent, the state of $\underline{i}$'s environment at time $\underline{t}$ can similarly be coded as the values of input variables for $\underline{i}$ at time $\underline{t}$ or, if preferred, at time $\underline{t} + \Delta$. In this fashion, most if not all principles of causal propagation can be captured by same-subject equations, even if that is not always the most insightful way to express them.

Thirdly, it is foundational in our causality intuitions that the cause/effect relation is transitive, irreflexive, and hence anti-symmetric--in short, a strict partial order. That is, if $\underset{\sim}{e}_i$, $\underset{\sim}{e}_j$, and $\underset{\sim}{e}_k$ are events such that $\underset{\sim}{e}_i$ is a cause of $\underset{\sim}{e}_j$, and $\underset{\sim}{e}_j$ is in turn a cause of $\underset{\sim}{e}_k$, then $\underset{\sim}{e}_i$ is thereby also mediately a cause of $\underset{\sim}{e}_k$. However, if $\underset{\sim}{e}_i$ is a cause of $\underset{\sim}{e}_j$, $\underset{\sim}{e}_j$ is not a cause of $\underset{\sim}{e}_i$, nor is any $\underset{\sim}{e}_i$ a cause of itself. We shall take transitivity to imply for the idealized extreme of single-input causality that if $\underset{\wedge}{y} = \phi(\underset{\wedge}{z})$ and $\underset{\wedge}{z} = \psi(x)$ are both strict causal regularities over population P, then so is their composition $\underset{\wedge}{y} = \phi\psi(x)$. Less simplistically, we presume that whenever $\underset{\wedge}{y} = \phi(\underset{\wedge}{Z},\underset{\wedge}{z}')$ and $\underset{\wedge}{z}' = \psi(\underset{\wedge}{X})$ are causal regularities, there is also a causal regularity $\underset{\wedge}{y} = \theta(\underset{\wedge}{W})$ in which $\underset{\wedge}{W}$ is a not-necessarily-proper subtuple of $\langle \underset{\wedge}{Z},\underset{\wedge}{X} \rangle$, even though the latter is not always simply $\underset{\wedge}{y} = \phi(\underset{\wedge}{Z},\psi(\underset{\wedge}{X}))$. (See p. 2.22ff. In fact, analysis of how causal regularities compose is the main concern of Chapter 2.) And we also assume that the strict partial order of causality among events is reflected by a strict partial order on the totality of variables that participate in same-subject regularities in any population P. That is, the (same-subject) causal-source relation in P is presumably transitive, anti-symmetric, and irreflexive.

Fourthly, we posit that when variables $\underset{\wedge}{X}$ are a strictly complete source of variable $\underset{\wedge}{y}$ in population P, the causal regularity $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ by which $\underset{\wedge}{X}$ determines $\underset{\wedge}{y}$ in P is unique even when, due to less-than-full dispersion of $\underset{\wedge}{X}$ in P, there exists a plurality of functions $\{\phi_k\}$ on $\underset{\wedge}{X}$'s logical range such that $\underset{\wedge}{y} = \phi_k(\underset{\wedge}{X})$ in P. (Variables $\underset{\wedge}{X} = \langle \underset{\wedge}{x}_1,\ldots,\underset{\wedge}{x}_m \rangle$ are "fully dispersed" in P just in case all logically possible tuples of values on $\underset{\wedge}{X}$ occur in P.) The argument for this is straightforward up to a point: If $\propto$ comprises the genuinely relevant boundary conditions in P under which a subject's value of $\underset{\wedge}{X}$ necessitates its having one value rather than another of $\underset{\wedge}{y}$, and P is contained in some larger population P* of subjects satisfying $\propto$ within $\underset{\wedge}{X}$ is fully dispersed, then $\underset{\wedge}{y} = \phi_k(\underset{\wedge}{X})$ in P* for only one function ϕ_k on $\underset{\wedge}{X}$'s logical range and this $\underset{\wedge}{y} = \phi_k(\underset{\wedge}{X})$ is then also the only dependency of $\underset{\wedge}{y}$ upon $\underset{\wedge}{X}$ in P*'s subpopulation P that qualifies as causal. Unhappily, unless each variable in $\underset{\wedge}{X}$ has only a small finite number of values, it is rather unlikely that any population of extant

even if full dispersion of $\underset{\wedge}{X}$ is not precluded by $\underset{\wedge}{X}$'s own causal origins.

subjects satisfying α is large enough to disperse $\underset{\wedge}{X}$ fully, But even then it seems proper to postulate that only one of the functions carrying $\underset{\wedge}{X}$ into $\underset{\wedge}{y}$ in $\underline{P}$ is truly causal. And if this causal equation is functionally irreducible and of a known (or assumed) restricted form Φ, it is uniquely identified by $\underset{\wedge}{y}$'s Φ-form regression upon $\underset{\wedge}{X}$ in $\underline{P}$ so long as $\underset{\wedge}{X}$'s dispersion in $\underline{P}$ is not strongly constrained in certain ways coordinate with form Φ.

Finally, we try to conceive of "causality" in a loose or generic sense that includes grades of molar determination that may well be just epiphenomenal abstractions from more genuine causal connections within ensembles of micro-events, yet which behave (or are rationally thought to behave) so much like basic causality, especially its partial-ordering of variables and the inference patterns it sustains, that for all practical purposes the difference is negligible. Just what is at issue here is difficult to make clear in a few words (or even in a great many of them), insomuch as the theory of molar causality is still in its prepartum infancy. But nearly any real-life causal story will serve to illustrate, such as the genetic example invoked earlier: When Jane's height at time $\underline{t}$ is attributed to the genetic constitution of her parents at the time of her conception, these properties of Jane's father and mother as momentary wholes are composites (not so?) of the chromatin characters of all their respective cells at that time, whereas roughly speaking it is only the chromatin in one particular cell from each parent that matters for Jane's later height. ("Roughly speaking" because Jane's genes may not be the only causal route by which her height is affected by her parents' biology.) Moreover, when we express this influence by saying that the height-of-individual-$\underline{i}$-at-time-$\underline{t}$ variable $(\underset{\wedge}{h})$ is affected by the variables $\langle \underset{\wedge}{g}_f, \underset{\wedge}{g}_m \rangle$ whose values for $\underline{i}$-at-$\underline{t}$ respectively correspond to certain genetic features of $\underline{i}$'s father and mother at $\underline{i}$'s conception, the attributes represented most directly by these variables' values are the existentially quantified relational properties here-and-now of having had a father and a mother of such-and-such genetic kinds, whereas the real causes of a particular $\underline{i}$'s height at time $\underline{t}$ in this fashion are the events consisting of certain specific individuals other than $\underline{i}$ being the way they were in the relevant respects at a

specific time other than $\underset{\wedge}{t}$. Still again, if $\underset{\wedge}{g}_0$ is a numerically scaled genetic variable such that $\underset{\wedge}{i}$'s values of $\underset{\wedge}{g}_f$ and $\underset{\wedge}{g}_m$ at $\underset{\wedge}{t}$ are respectively defined to be the values of $\underset{\wedge}{g}_0$ for $\underset{\wedge}{i}$'s father and mother at $\underset{\wedge}{i}$'s conception, and $\underset{\wedge}{g}_f$ and $\underset{\wedge}{g}_m$ contribute equally and linearly to $\underset{\wedge}{h}$, we may say simply that the main genetic cause of $\underset{\wedge}{i}$'s height at $\underset{\wedge}{t}$ is $\underset{\wedge}{i}$'s mean-parental-$\underset{\wedge}{g}_0$-value $(\underset{\wedge}{g}^*)_{\text{even}}$ though there is no genuine causal mediation of $\langle \underset{\wedge}{g}_f, \underset{\wedge}{g}_m \rangle \rightarrow \underset{\wedge}{h}$ by variable $\underset{\wedge}{g}^* =_{\text{def}} (\underset{\wedge}{g}_f + \underset{\wedge}{g}_m)/2$.

This last example instantiates the only tentative principle of molar causality that will be suggested here: If the causal regularity by which variables $\langle \underset{\wedge}{X}, \underset{\wedge}{Z} \rangle$ determine variable $\underset{\wedge}{y}$ (in $\underline{P}$) takes form $\underset{\wedge}{y} = \phi(\underset{\wedge}{X}, \theta_1(\underset{\wedge}{Z}))$, with tuples $\underset{\wedge}{X}$ and $\underset{\wedge}{Z}$ disjoint, we can abstract molar variable $\underset{\wedge}{z}^*$ as composite $\underset{\wedge}{z}^* =_{\text{def}} \theta_1(\underset{\wedge}{Z})$ of the variables in tuple $\underset{\wedge}{Z}$ and take $\underset{\wedge}{y} = \phi(\underset{\wedge}{X}, \underset{\wedge}{z}^*)$ to be a molar causal regularity in which $\underset{\wedge}{Z}$ is a (quasi)-causal source of $\underset{\wedge}{z}^*$ while $\underset{\wedge}{z}^*$ in turn mediates (quasi)-causally between $\underset{\wedge}{Z}$ and $\underset{\wedge}{y}$. To account for $\underset{\wedge}{z}^*$, we can cite $\underset{\wedge}{Z}$'s determination of $\underset{\wedge}{z}^*$ by molar/molecular abstraction. But also, more informatively, there may be molecular/molecular regularities $\underset{\wedge}{Z} = \psi_1(\underset{\wedge}{F})$ in which each variable in $\underset{\wedge}{Z}$ is causally determined by variables $\underset{\wedge}{F}$. Then $\underset{\wedge}{z}^* = \theta_1(\psi_1(\underset{\wedge}{F}))$ is a (quasi)-causal molar/molecular regularity whose mapping from $\underset{\wedge}{F}$ to $\underset{\wedge}{z}^*$ can generally be decomposed in many different ways $\theta_1(\psi_1(\_\_)) = \theta_2(\psi_2(\_\_)$ wherein ψ_2 is in general vector-valued. If we stipulate $\underset{\wedge}{F}^* =_{\text{def}} \psi_2(\underset{\wedge}{F})$ for some favored choice of ψ_2 (or more realistically, if our burgeoning theory of the phenomenon at issue gives us a concept of variables $\underset{\wedge}{F}^*$ that we later discover can best be viewed as analytic abstractions from a more complex array $\underset{\wedge}{F}$ of molecular variables), we can replace $\underset{\wedge}{z}^* = \theta_1(\psi_1(\underset{\wedge}{F})) = \theta_2(\psi_2(\underset{\wedge}{F}))$ by the molar/molar (quasi)-structural equation $\underset{\wedge}{z}^* = \theta_2(\underset{\wedge}{F}^*)$ under which $\underset{\wedge}{F}^*$ (quasi)-causally determines $\underset{\wedge}{z}^*$. Continuing in this fashion, felicitous selection of $\underset{\wedge}{z}^*$, $\underset{\wedge}{F}^*$, and other compositing abstractions including in all likelihood $\underset{\wedge}{y}$ to start with, may construct a partial order of molar variables linked by law-like dependencies that are simplified images of genuinely causal but vastly more complicated molecular regularities whose explicit details generally exceed our powers to comprehend or even discover. (It is important to appreciate in this respect that our development of concepts at various levels of the molar/

molecular hierarchy is generally top-down rather than bottom-up.) How to tell which systems of molar variables best summarize reality's hard-core causal structure, and to what degree the alternatives may be arbitrary, like choices among scaling units, are matters on which no reasoned theory has yet appeared in the multivariate (or for that matter philosophical) literature, even though multivariate practice has already engaged the issue in disputes over optimalities of axis placement in spaces of variables. Fortunately, MODA is not committed to any specific position on such matters. But it does openly acknowledge that applied data analysis is, wittingly or unwittingly, profoundly concerned with them.

Linear causality.

For mathematical convenience--so powerful a convenience that we seldom have much practical alternative--we largely restrict quantitative models of causality in applied data analysis to cases in which a variable $\underset{\wedge}{y}$ is determined in $\underline{P}$ jointly by variables $\underset{\wedge}{X} = \langle \underset{\wedge}{x}_1, \ldots, \underset{\wedge}{x}_n \rangle$ and $\underset{\wedge}{Z} = \langle \underset{\wedge}{z}_1, \ldots, \underset{\wedge}{z}_r \rangle$ in accord with some linear structural equation

$$(1) \qquad \underset{\wedge}{y} = \underline{w}_0 + \sum_{i=1}^{n} \underline{w}_i \underset{\wedge}{x}_i + \sum_{j=1}^{r} \underline{v}_j \underset{\wedge}{z}_j = \underline{w}_0 + \underset{1 \times n}{\underline{w}} \underset{\wedge}{X} + \underset{1 \times r}{\underline{v}} \underset{\wedge}{Z} \qquad \text{(assumed)},$$

where $\underset{1 \times n}{\underline{w}} = \langle \underline{w}_1, \ldots, \underline{w}_n \rangle$ and $\underset{1 \times r}{\underline{v}} = \langle \underline{v}_1, \ldots, \underline{v}_r \rangle$ are row vectors of conjoint causal weights. (Later, $\underset{\wedge}{X}$ will comprise observed variables while variables $\underset{\wedge}{Z}$ are unobserved; but for now, the partition between $\underset{\wedge}{X}$ and $\underset{\wedge}{Z}$ in (1) is arbitrary.) Whenever $\underset{\wedge}{y}$'s dependence on $\langle \underset{\wedge}{X}, \underset{\wedge}{Z} \rangle$ in $\underline{P}$ is characterized by an equation of form (1), we shall say that $\underset{\wedge}{y}$ is LE-determined (i.e., Linearly and Errorlessly) by $\langle \underset{\wedge}{X}, \underset{\wedge}{Z} \rangle$ in $\underline{P}$.

Though admittedly an idealization, equation (1) is not nearly so unrealistic as the several ways in which it is prima facie restrictive may make it appear. Since we shall be making much of LE-determination later, it is thus wise to begin with some clarification of its methodological status, especially its cogency as a model of real-world causality.

In the first place, without any loss of linearity's formal power, addition and multiplication in (1) can be interpreted as any mathematical operations having

the combinatorial properties that constitute an abstract-algebraic ring, so long as this construal of the operators is fixed through the full array of form-(1) equations under consideration. This is equivalent to saying that even when plus and times in (1) are defined by ordinary arithmetic, we are free to choose whatever numerical scales for our variables most closely linearize their structural relations. How to identify linearity-wise optimal scales for our data variables in practice is a very nice question indeed (For hidden sources, this scaling problem does not arise.) But the theory of multivariate relations presupposes the existence of many things to which we have only imperfect operational access, and ideal scales can just as well be included among them.

Secondly, if $\hat{y}$'s causal determination in P is only stochastically lawful, with just the expectation of a conditional probability distribution for $\hat{y}$ strictly determined by antecedent causes, we can formally treat $\hat{y}$'s divergence from that expectation as an additional "source" of $\hat{y}$ and thereby regain the mathematics of strict determination. That is, if needed we let one of variables $\hat{z}_j$ in (1) be whatever in $\hat{y}$ is irreducibly indeterministic. To be sure, even if $\hat{y}$'s less-than-perfect determination by the totality $\hat{S}$ of its real sources can adequately be expressed by a function ϕ under which $\hat{S}$ causally imposes a tendency on $\hat{y}$ to take value $\phi(\hat{S})$, it is still moot whether $\hat{y}$-tendency $\phi(\hat{S})$ is best construed as $\hat{y}$'s statistical expectation. But it is extraordinarily difficult to conceive how $\hat{S}$ might cause $\hat{y}$ only semi-deterministically except by envisioning a residual $\hat{e}_y$ such that $\hat{y} = \phi^*(\hat{S},\hat{e}_y)$ for some function ϕ^* that is tantamount to a determinstic causal function under which residual $\hat{e}_y$ behaves as though it is a source of $\hat{y}$ that has no causes of its own. Lacking any other way to conceive of partial determinacy, it is not merely appropriate but operationally necessary to treat causality as locally deterministic: Not all variables need have causes, but those having any at all are determined completely by their sources.

Thirdly, if the number of variables $\{z_k\}$ needed to supplement $\underset{\wedge}{X}$ into a strictly complete source of y is infinite--as in fact may well be the case--we can reasonably presume that there is a sufficiently large finite tuple $\underset{\wedge}{Z}$ of these supplementary sources such that the joint determination of y by $\underset{\wedge}{X}$ and $\{z_k\}$ can be expressed by a structural equation $y = \phi(\underset{\wedge}{X}, \underset{\wedge}{Z}, z^*)$ in which z^* is some composite of the variables in $\{z_k\}$ other than in $\underset{\wedge}{Z}$ and which either contributes negligibly to $\phi(\underset{\wedge}{X}, \underset{\wedge}{Z}, z^*)$ or for practical purposes behaves like a stochastic residual. This assumption rests upon a certain amount of blind faith; but the mathematics of laws containing an infinite number of independent variables is not only disturbingly enigmatic but to my knowledge has not been seriously explored apart from my own tortuous effort in Rozeboom (1978).

Fourth, since we do not require all of coefficients $w_1, \ldots, w_n$ in (1) to be non-zero, we can allow some or all of variables $\underset{\wedge}{X}$ (and similarly for $\underset{\wedge}{Z}$) to be only putative sources of y, conjoint with the others, that are not truly so. For if any of the x_i do not in fact share joint responsibility for y with $\underset{\wedge}{Z}$ and the rest of $\underset{\wedge}{X}$, that is expressed by zero values of the corresponding w_i. It is precisely to allow zero structural weights that we have provided (p. 10a above) for a sense of joint causation under which a structural equation's independent variables need not be strictly joint sources of its dependent variable. We do not want to admit irrelevant variables as causal antecedents so promiscuously that causal determination loses its partial-order character. But it is harmlessly convenient to allow that if $y = \phi(\underset{\wedge}{X})$ is a structural equation for the complete joint determination of y by variables $\underset{\wedge}{X}$ in population P, and z is any variable that does not mediate any of $\underset{\wedge}{X}$'s causal influence upon y, then apart from possible exclusions not yet motivated,

$$y = \phi(\underset{\wedge}{X}) + 0 \cdot z$$

is also a structural equation for joint determination of y by $\langle \underset{\wedge}{X}, z \rangle$ in P. This definition is recursive on the base of structural equations that express strict causal regularities (i.e., that contain no independent variables that do not work jointly with the rest), and implies that not all independent variables in a

structural equation necessarily count as sources of its dependent variable.

Fifth, and of great importance, the formal linearity of (1) does not preclude any x_i in X (or z_j in Z) actually affecting y in much more curvilinear fashion. For we allow that some of the other variables in X and Z may be fixed nonlinear functions of x_i and perhaps other variables, or that x_i and certain other variables in $\langle X, Z \rangle$ may be different nonlinear abstractions from one or more "real" variables not separately included in $\langle X, Z \rangle$. For example, if y is a quadratic function of just two variables x_1 and x_2, the dependency has form

$$y = w_0 + w_1 x_1 + w_2 x_2 + v_1 x_1^2 + v_2 x_2^2 + v_3 (x_1 x_2) ,$$

which can be subsumed under (1) by taking, say, $X = \langle x_1, x_2 \rangle$ and $Z = \langle z_1, z_2, z_3 \rangle = \langle x_1^2, x_2^2, x_1 x_2 \rangle$. Again, if y is determined nonlinearly by a categorical variable x^* (or by the cartesian product of a tuple of categorical variables) which has just s alternative values ("levels"), we can define a tuple $X^* = \langle x_1^*, \ldots, x_{s-1}^* \rangle$ of binary variables such that x_k^* ($k = 1, \ldots, s-1$) takes value 1 or 0 according to whether the subject is or is not at the kth level on x^*, and have that the linear regression of y upon X^* is identical with y's unrestricted curvilinear regression on x^*. (This technique, with a special rotation of X^* to align with "main effects" and various orders of "interaction," is how ANOVA/ANCOVA subsumes y's dependency on multiple categorical variables under the general linear model.) So long as we do not require all variables in $\langle X, Z \rangle$ to be logically independent of one another, the L in LE-determination is not an essential loss of generality. Nonlinearities do create applied problems, especially when the data are insufficiently abundant to permit recovery of a great many parameters; and parameterizing curvilinear functions linearly does

not avoid certain obscurities of interpretation to be reluctantly noted as we proceed. But that is no objection to the general linear model's cogency for theoretical analysis.

Finally, it is of the utmost importance to be clear that regardless of any form idealizations imposed on causal regularities, the law of any variable's causal determination in $\underline{P}$ is profoundly nonunique. It is flagrantly not the case that output $\underset{\wedge}{y}$ has only one complete joint source in $\underline{P}$, not even if we exclude joint sources of $\underset{\wedge}{y}$ that are not strictly so. For if $\langle \underset{\wedge}{X}, \underset{\wedge}{Z} \rangle$ determines $\underset{\wedge}{y}$, the tuple of variables derived from $\langle \underset{\wedge}{X}, \underset{\wedge}{Z} \rangle$ by replacing any variable therein by a complete joint source of its own is also a complete joint source of $\underset{\wedge}{y}$. For example, if

$$\underset{\wedge}{y} = \underline{w}_1 \underset{\wedge}{z}_1 + \underline{w}_2 \underset{\wedge}{z}_2 , \quad \underset{\wedge}{z}_1 = \underline{v}_1 \underset{\wedge}{x}_1 + \underline{v}_2 \underset{\wedge}{x}_2 , \quad \underset{\wedge}{z}_2 = \underline{v}_3 \underset{\wedge}{x}_1 + \underline{v}_4 \underset{\wedge}{x}_3$$

are all causal laws in $\underline{P}$, where z_2 does not mediate between $\langle \underset{\wedge}{x}_1, \underset{\wedge}{x}_2 \rangle$ and $\underset{\wedge}{z}_1$ nor $\underset{\wedge}{z}_1$ between $\langle \underset{\wedge}{x}_1, \underset{\wedge}{x}_3 \rangle$ and $\underset{\wedge}{z}_2$, then variables $\langle \underset{\wedge}{x}_1, \underset{\wedge}{x}_2, \underset{\wedge}{x}_3, \underset{\wedge}{z}_1, \underset{\wedge}{z}_2 \rangle$ are all causal sources of $\underset{\wedge}{y}$ in $\underline{P}$ (as are also in turn all causal sources of the $\underset{\wedge}{x}_i$), and each of

$$(2\underline{a}) \qquad \underset{\wedge}{y} = \underline{w}_1 \underset{\wedge}{z}_1 + \underline{w}_2 \underset{\wedge}{z}_2 + 0 \cdot \underset{\wedge}{x}_1 + 0 \cdot \underset{\wedge}{x}_2 + 0 \cdot \underset{\wedge}{x}_3$$

$$(2\underline{b}) \qquad \underset{\wedge}{y}' = \underline{w}_1 \underset{\wedge}{z}_1 + (\underline{w}_2 \underline{v}_3) \underset{\wedge}{x}_1 + 0 \cdot \underset{\wedge}{x}_2 + (\underline{w}_2 \underline{v}_4) \underset{\wedge}{x}_3$$

$$(2\underline{c}) \qquad \underset{\wedge}{y} = \underline{w}_2 \underset{\wedge}{z}_2 + (\underline{w}_1 \underline{v}_1) \underset{\wedge}{x}_1 + (\underline{w}_1 \underline{v}_2) \underset{\wedge}{x}_2 + 0 \cdot \underset{\wedge}{x}_3$$

$$(2\underline{d}) \qquad \underset{\wedge}{y} = (\underline{w}_1 \underline{v}_1 + \underline{w}_2 \underline{v}_3) \underset{\wedge}{x}_1 + (\underline{w}_1 \underline{v}_2) \underset{\wedge}{x}_2 + (\underline{w}_2 \underline{v}_4) \underset{\wedge}{x}_3$$

is a form-(1) structural equation for $\underset{\wedge}{y}$'s production from its sources in $\underline{P}$. (For simplicity, these hypothesized determinations omit residual sources that in practice we not only expect but, to ward off demons of multicollinearity, actively desire so long as they are orthogonal to the sources explicitly acknowledged.) It would be a monumental blunder to think that (2b)-(2d) are less genuine, less real, or less truly causal than (2a) because they do not give $\underset{\wedge}{y}$'s dependence upon its immediate causes. So far as we have any reason to believe, the effect of any source variable $\underset{\wedge}{x}_i$ on any output variable $\underset{\wedge}{y}$ is always mediated by some tuple $\underset{\wedge}{F}$ of intervening source variables such that $\underset{\wedge}{x}_i$'s (partial) causation of $\underset{\wedge}{y}$ is by virtue of $\underset{\wedge}{x}_i$'s affecting $\underset{\wedge}{F}$ and $\underset{\wedge}{F}$'s affecting $\underset{\wedge}{y}$. Consequently, when variables $\underset{\wedge}{X}$ determine $\underset{\wedge}{y}$ partially or even completely in $\underline{P}$, it is not meaningful to ask what is the law by which $\underset{\wedge}{X}$ brings about

$\hat{y}$ in $\underline{P}$. Rather, this question has a unique answer only relative to some choice of supplementary variables $\hat{Z}$ such that $\hat{X}$ and $\hat{Z}$ together include a strictly complete source of $\hat{y}$ in $\underline{P}$ (with any ultimate stochastic uncertainty in $\hat{y}$ as an additional "source" in $\hat{Z}$). Relative to $\hat{Z}$, there will be some structural equation $\hat{y} = \phi(\hat{X},\hat{Z})$ characterizing $\hat{y}$'s dependence on $\langle\hat{X},\hat{Z}\rangle$ in $\underline{P}$; and if this equation is linear, we can further represent the $\hat{y}$-influence of each $\hat{x}_i$ in $\hat{X}$, conjoint with and relative to the remainder of $\langle\hat{X},\hat{Z}\rangle$, by a single coefficient. (If some variables in $\langle\hat{X},\hat{Z}\rangle$ are nonlinear abstractions from others, however, we must be careful how we interpret these coefficients. In particular, the coefficient of $\hat{x}_i$ cannot then be construed to tell how changes in $\hat{x}_i$ affect $\hat{z}$ when the other variables in $\langle\hat{X},\hat{Z}\rangle$ are held constant.) But if $\langle\hat{X},\hat{Z}_1\rangle$ and $\langle\hat{X},\hat{Z}_2\rangle$ each LE-determine $\hat{y}$ in $\underline{P}$, the coefficients of $\hat{X}$ for $\hat{y}$ are not the same relative to $\hat{Z}_1$ as they are relative to $\hat{Z}_2$ except under special circumstances noted later.

The nature of this relativity is entirely straightforward: Roughly speaking, the coefficients in (1) express what each $\hat{x}_i$ and $\hat{z}_j$ in $\langle\hat{X},\hat{Z}\rangle$ contributes to $\hat{y}$ independently of the other variables in this particular tuple of $\hat{y}$-sources. So to the extent that $\hat{x}_i$'s causal import for $\hat{y}$ is mediated by $\hat{z}_j$, inclusion of $\hat{z}_j$ in $\hat{Z}$ withdraws some of the weight for $\hat{y}$ that $\hat{x}_i$ would receive if $\hat{Z}$ were not chosen to intercept the $\hat{x}_i \rightarrow \hat{y}$ connection, and assigns this instead to $\hat{z}_j$ (or perhaps to other variables in $\hat{Z}$ that in turn mediate between $\hat{z}_j$ and $\hat{y}$). Thus in example (2), variables $\langle\hat{x}_1,\hat{x}_2,\hat{x}_3\rangle$ are a complete joint source of $\hat{y}$ whose structural equation (2d) for determining $\hat{y}$ could be identified by computing the regression of $\hat{y}$ upon just $\langle\hat{x}_1,\hat{x}_2,\hat{x}_3\rangle$. But if $\hat{y}$ is regressed on all of $\langle\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{z}_1,\hat{z}_2\rangle$ (after adding vanishingly small residuals to the determinations of $\hat{z}_1$ and $\hat{z}_2$ by $\langle\hat{x}_1,\hat{x}_2,\hat{x}_3\rangle$ to break the multicollinearities), we instead obtain equation (2a) in which only $\hat{z}_1$ and $\hat{z}_2$ have nonzero coefficients because the effects of $\langle\hat{x}_1,\hat{x}_2,\hat{x}_3\rangle$ on $\hat{y}$ are wholly mediated by $\langle\hat{z}_1,\hat{z}_2\rangle$. And when $\hat{y}$ is regressed just on $\langle\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{z}_1\rangle$ or just on $\langle\hat{x}_1,\hat{x}_2,\hat{x}_3,\hat{z}_2\rangle$, thereby recovering (2b) or (2c) respectively, $\hat{x}_1$ retains nonzero coefficient (albeit a different one in the two cases) because $\hat{x}_1$ affects $\hat{y}$ through multiple lines of connection not all of which are intercepted by mediators conjoined with $\hat{x}_1$

in the $\underset{\wedge}{y}$-sources to which these coefficients are relative.

Terminological note: When $\underset{\wedge}{X}$ and $\underset{\wedge}{Z}$ are jointly a complete source of y, we shall speak of $\underset{\wedge}{X}$'s coefficients in structural equation (1) as relative either to $\underset{\wedge}{Z}$ or to $\langle \underset{\wedge}{X}, \underset{\wedge}{Z} \rangle$, whichever seems more natural in context.

"Regularities" and their transducers.

Present usage of schema "$\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$" or more explicitly "$\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ in $\underline{P}$" to denote regularities, causal or otherwise, in population $\underline{P}$ requires a special understanding if it is to do its job properly. When ϕ and ψ are both functions from the space of all logically possible values on variables $\underset{\wedge}{X}$ into values of variable y but are not set-theoretically the same function from this domain into this range--i.e. when $\phi(\underline{X}) \neq \psi(\underline{X})$ for at least one $\underset{\wedge}{X}$-value $\underline{X}$--we want to be able to say that hypothesized regularities $y = \phi(X)$ and $y = \psi(X)$ in $\underline{P}$ are different regularities even when $\phi(\underline{X}) = \psi(\underline{X})$ for every value $\underline{X}$ of $\underset{\wedge}{X}$ that actually occurs in $\underline{P}$. (Need for this distinction arises when $\underset{\wedge}{X}$ is not fully dispersed in $\underline{P}$.) However, what "$\phi(\underset{\wedge}{X})$ (in $\underline{P}$)" literally refers to is the composition into ϕ of the restriction $\underset{\wedge}{X}_P$ of variables $\underset{\wedge}{X}$ to population $\underline{P}$, i.e. $\underset{\wedge}{X}_P$ is the function mapping members of $\underline{P}$ into their $\underset{\wedge}{X}$-values. So if "regularity $\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$ (in $\underline{P}$)" were understood to designate merely the hypothesized fact that the restriction $\underset{\wedge}{y}_P$ of y to $\underline{P}$ is identical with the composition of $\underset{\wedge}{X}_P$ into ϕ, i.e. that $\underset{\wedge}{y}_P = \phi\underset{\wedge}{X}_P$, it would have the same referent as "regularity $\underset{\wedge}{y} = \psi(\underset{\wedge}{X})$ (in $\underline{P}$)" whenever ϕ and ψ are identical just over the values of $\underset{\wedge}{X}$ that actually occur in $\underline{P}$.

Accordingly, when we write "$\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$" to refer to a hypothesized regularity in an implicitly specified population $\underline{P}$, we shall understand this to refer to a 2-tuple whose first component is the (hypothesized) extensional generality that function $\phi(\underset{\wedge}{X})$ is identical with function $\underset{\wedge}{y}$ over $\underline{P}$, i.e. the fact that $\underline{y}_s = \phi(\underline{X}_s)$ for all subjects $\underline{s}$ in $\underline{P}$, and whose second component is what we will call the regularity's transducer, namely, the full function ϕ named in "$\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$". Then if $\phi \neq \psi$, terms "$\underset{\wedge}{y} = \phi(\underset{\wedge}{X})$" and "$\underset{\wedge}{y} = \psi(\underset{\wedge}{X})$" designate different regularities in $\underline{P}$ even when

$\phi\hat{X}_P = \hat{y}_P = \phi\hat{X}_P$ because their transducers are different; and we can hence claim of one that it is causal without entailing that the other is causal as well. Specifically, when we conjecture that $\hat{y} = \phi(\hat{X})$ is a causal regularity in $\underline{P}$, we envision that there is some attribute (or ensemble of attributes) $\propto$ common to members of $\underline{P}$, and a principle of natural necessitation, such that for every possible value $\underline{X}$ of $\hat{X}$, joint possession of properties $\propto$ and $\underline{X}$ necessitates an accompanying value $\phi(\underline{X})$ of $\hat{y}$ as well, regardless of whether $\underline{X}$ is actually instantiated in $\underline{P}$ or elsewhere. In this way, attributing causality to a regularity makes essential reference to the regularity's complete transducer, and explains the force of counterfactual statements such as "Although no member of $\underline{P}$ in fact has value $\underline{X}^*$ of $\hat{X}$, if any subject $\underline{s}$ in $\underline{P}$ were to have $\hat{X}$-value $\underline{X}^*$, $\underline{s}$'s value of $\hat{y}$ would be $\phi(\underline{X}^*)$."