Cowles Commission Discussion Paper: Statistics: No. 340

November 23, 1949

A Note on Models and Identification

by

Carl Christ

A theory of the behavior through time of a set of economic variables $y_1, \ldots, y_G$ can be expressed in the following general form:

$$(1) \qquad T \begin{pmatrix} y_1 \\ \vdots \\ y_G \\ z_1 \\ \vdots \\ z_K \end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_G \end{pmatrix}$$

where T is a transformation, $z_1, \ldots, z_K$ are predetermined variables representing previous conditions or exogenous influences which do not need to be explained by the theory, and $u_1, \ldots, u_G$ are random disturbances having a joint probability distribution Ψ. The y_g are known as jointly dependent variables, $g = 1, \ldots, G$.

A structure is defined as a pair (T_o, Ψ_o), where T_o is a completely specified transformation T and Ψ_o is a completely specified joint distribution Ψ. For example, in a simple case where $G = 1$ and $K = 0$, the pair

$$(2) \qquad \begin{aligned} &T: && y_1 + \alpha = u_1 \\ &\Psi: && \Psi(u_1) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{u_1} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \end{aligned}$$

is not a structure because certain parameters α, μ and σ are not specified; if we assign specific values to them, such as $\alpha = 3.87$, $\mu = 0$, and $\sigma = 1.26$, then (2) becomes a structure.

Another example of a structure with $G = 1$ and $K = 0$ is given by the following graphs, representing a completely specified pair (T_o, Ψ_o):

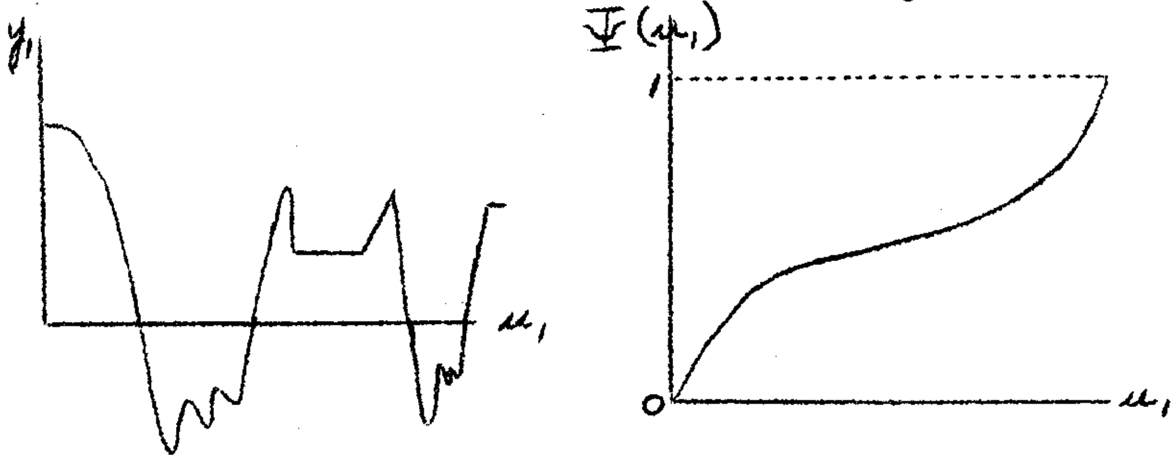

Figure 1.

Each structure expresses a completely specified hypothesis about the behavior of the set of economic variables y_g. Consider the set C of all conceivable structures; it is an infinite set. If we state no a priori restrictions which we believe the true structure must satisfy, then any member of C is a priori admissible as a hypothesis. If we do state certain restrictions upon the structure, then some of the members of C are no longer admissible a priori because they violate the restrictions. We will define a model M as a set of a priori admissible structures, i.e., the subset of C whose members do not violate the given a priori restrictions (if any) upon the structure.

The set C of all conceivable structures can be divided into an infinity of infinite subsets C_i, each subset C_i containing exactly those structures which are compatible with some conceivable state of nature, i.e., with some probability distribution of the observed values of the variables y_g and z_k, $g = 1, \ldots, G$ and $k = 1, \ldots, K$. It is clear that there is an infinity of the C_i because the number of conceivable states of nature ~~structures~~ is infinite, and that each C_i is infinite because any state of nature can be explained by an infinity of hypotheses. Every structure of C belongs to at least one C_i; the C_i need not be disjunct because there are hypotheses which will explain everything (these are meaningless, of course).

The structures which are left to us as possible hypotheses after we have placed our a priori restrictions (if any) on the set C of conceivable structures, thus obtaining our model M, and after we have observed a certain state of nature, say the j^{th}, are those belonging to the intersection of M and C_j, where C_j corresponds to the observed j^{th} state of nature. Let us call this intersection A_j. Then $A_j = M \cap C_j$.

If A_j is empty for some j, say j_o, we have placed too severe a priori restrictions on the structure -- we have no possible hypothesis in case the ${j_o}^{th}$ state of nature materializes. If A_j has two or more members for some j, say j_2, then we have not restricted the structure enough -- we do not know which hypothesis to choose in case the ${j_2}^{th}$ state of nature materializes. If A_j has exactly one member for some j, say j_1, then our hypothesis is uniquely determined if the ${j_1}^{th}$ state of nature materializes. We say a structure is identified within a model M if it belongs to a set A_j which has exactly one member. We say an identified structure is overidentified if it remains identified when some of the a priori restrictions are removed; a set of restrictions which can be so removed is called a set of overidentifying restrictions. We would like to choose the model M in such a way that to each state of nature there corresponds exactly one structure (hypothesis), or equivalently, so that each structure in the model is identified.

A simple example of a model is provided by equations (2) above. If μ is known, all structures belonging to this model are identified. If μ is not known, none of its structures is identified.

Another example of a model, again with $G = 1$ and $K = 0$, is the following, where $\bar{y}_1$ is the mean of a set of observed values of y_1:

$$
(3) \qquad
\begin{aligned}
T:&\ \begin{cases} y_1 \text{ is ordained by Fate if } \bar{y}_1 = 0 \\ \qquad y_1 = u_1 \qquad\qquad \text{if } \bar{y}_1 < 0 \\ y_1 - \bar{y}_1 = u_2 \qquad\qquad \text{if } \bar{y}_1 > 0 \end{cases} \\
\Psi:&\ \begin{cases} \Psi(u_1) \text{ is the distribution } \Psi_0 \text{ shown in Figure 1.} \\ \Psi(u_2) = \dfrac{1}{\sqrt{2\pi}\,\sigma} \displaystyle\int_{-\infty}^{u_1} e^{-\frac{x^2}{2\sigma^2}}\, dx \end{cases}
\end{aligned}
$$

This is a perfectly permissible model, though it has little to make it plausible. A T of this kind (minus the mystical touch) might be used, however, with zero-mean disturbances, in case one knew for sure that y_1 was a random variable with positive mean.

Practical work with a model is much more convenient if the model is a parametric family of structures and has a simple general form, preferably linear in variables but at least linear in parameters. Models of any desired complexity can be approximated by such a parametric family, with the help of polynomial expansions and restrictions on subsets of the parameters. Accordingly we shall consider only models which are parametric families of structures, linear in the parameters. Such a model looks like this:

$$
(4) \qquad \begin{aligned} &T: && \beta y' + \Gamma z' = u' \\ &\Psi: && \Psi(u) = \text{some explicit parametric function of } u \end{aligned}
$$

where β and Γ are matrices of parameters (both having G rows), and y', z', and u' are column vectors of jointly dependent, predetermined, and random variables, respectively. If the equations are linear in variables as well as in parameters, then β is square G x G, and y' and u' have G elements each. If the equations are not linear in variables but only in parameters, then β is not square but has more than G columns; u' still has G elements, however. In this latter case, some of the elements of y may be functions of other elements of y, or functions of other elements of y and elements of z, and similarly some of the elements of z may be functions of other elements of z, provided, these functions have no unknown parameters. E.g., we may have

$y_1 = y_1$	$z_1 = z_1$
$y_2 = y_2$	$z_2 = z_2$
$y_3 = y_3$	$z_3 = z_3$
$y_4 = y_1 y_2$	$z_4 = z_2 z_3$
$y_5 = y_4 - .001 z_1$	$z_5 = z_2^2$
$y_6 = (y_4 - .001 z_1)/y_3$	etc.
etc.	

but not $y_5 = y_4 - \alpha z_1$

Formal conditions for the identification of structures within such a model have been developed. We will consider here only models whose structures are all identified.

The state of nature (i.e., the probability distribution of the variables) is not observed exactly; only a sample from it is observed. Therefore the values of the parameters in the structure cannot be found exactly; they may be consistently estimated via the maximization of their likelihood function. This of course assumes that the model includes the true structure among its members. The efficiency of the estimates decreases if less a priori information in the form of restrictions on the structure is used, and becomes zero if the restrictions are reduced until the structure is no longer identified.