(Draft Copy)

On the Statistical Interpretation of Structural Equations

Judea Pearl
Cognitive Systems Laboratory
Computer Science Department
University of California, Los Angeles, CA 90024
judea@cs.ucla.edu

1 Background

In a recent discussion [8], Arthur S. Goldberger gives the following interpretation to structural equations:

> an economist might arrive at a model like this:
>
> $$y_1 = a_1 y_2 + a_2 x_1 + u_1 \quad (1a)$$
> $$y_2 = a_3 y_1 + a_4 x_2 + u_2 \quad (1b)$$
>
> Here, y_1=quantity, y_2=price, x_1=income, x_2=wage rate, u_1=demand shock, $u_2 =$ supply shock. The first equation states that household demand depends on price and household income (which are observable) and an unobserved factor. The systematic part of this equation, $a_1 y_2 + a_2 x_1$, may be interpreted as the expected quantity demanded if y_2 and x_1 were fixed. So the a_1 and a_2 parameters have natural meaning for the economist. Similarly for the second, supply, equation. The exclusion of x_2 from the demand equation and of x_1 from the supply equation reflect the economist's understanding of household and producer behavior.

In a reply to Goldberger, Nanny Wermuth [23] notes that if one seriously calculates "the expected quantity demanded if y_2 and x_1 were fixed" using the model described in (1), the result does not match the interpretation advanced by Goldberger. Specifically, assuming u_1 and u_2 are zero-mean disturbances (independent on X_1 and X_2), Wermuth finds

$$E(Y_1 \mid Y_2 = y_2, X_1 = x_1) \neq a_1 y_2 + a_2 x_1$$

(unless further assumptions are made) and concludes that "the parameters in (1) cannot have the meaning Arthur Goldberger claims they have."[1]

This exchange between a statistician and an economist exemplifies the long history of tension between regression analysis and structural equations modeling, which dates back to the inception of the latter by Wright [27], Tinbergen [22], and Haavelmo [10]. Through the works of Blalock [2], Duncan [5], Simon [20], and Goldberger [9], structural equations (or *path analysis*) methodology has gained acceptance and is currently used by most social scientists, econometricians, and psychologists [11, 1]. Nonetheless, statisticians have relentlessly criticized this methodology for the maverick way in which it integrates substantive knowledge with data analysis [16, 12, 14, 4, 6].

Economists prefer structural equations to regression models because the equations mirror the organization of substantive economic theories and, hence, structural parameters seem more natural and more fundamental than regression parameters. For example, Goldberger [7] states:

> the search for structural parameters is a search for invariant features of the mechanisms that generate observable variables. Invariant features are those which remain stable – or vary individually – over the set of populations in which we are interested. When regression parameters have this invariance, they are proper objects of research, and regression is an appropriate tool. But when, as appears to be the case in many social science areas, regression parameters lack this invariance, the proper objects of research are more fundamental parameters; and statistical tools which go beyond conventional regression are required.

Yet the difficulty of providing a consistent statistical interpretation for the parameters in structural equations, and especially the lack of an operational meaning for the zero coefficient terms, has been a constant source of embarrassment to economists. Whittaker [25], for example, remarks:

> The structural equation formulation may generate confusion for several reasons:
>
> 1. While formulating equations focuses on interactions present in the model, it does not always make clear the conditioning sets for those interactions missing from the model.
> 2. It is not always clear whether a parameter is free and to be estimated, or redundant and to be derived from the fitted parameters, or constrained, for instance, to be zero.

Wermuth and Cox [24] express similar sentiments: "in general, a zero coefficient in a structural equation does not correspond to an independence relation" and "for linear structural equations in general, the interpretation of equation parameters, be they present or missing, has to be derived from scratch for each model considered."

Indeed, if structural parameters were truly as fundamental and invariant as economists believe they are, then surely they should manifest themselves in some experimental setting,

[1]This inequality is known as "bias" in the structural equations literature, and is characteristic of nonrecursive systems. Goldberger is unquestionably aware of this inequality. An explanation of the discrepancy between Goldberger's and Wermuth's interpretations is offered in Section 2.2.

however hypothetical, and then we should expect economists to define these parameters by referring to the outcome of that experimental setting. Goldberger's interpretation lacks these features; his definition of $a_1 y_2 + a_2 x_1$ as "the expected quantity demanded if y_2 and x_1 were fixed" seems to clash with the calculus of expectation, while his justification for the exclusion of x_1 and x_2 as reflecting "the economist's understanding," is far from the empirically grounded justifications (say, in terms of conditional independence) statisticians have grown accustomed to. These difficulties have led generations of statisticians to doubt the empirical content of the structural equations enterprise and to ask repeatedly whether the enterprise is grounded in experience or sheerly metaphysical.

2 Reconciliation

In this note, I propose a clear statistical interpretation to structural equations, thus reconciling some of the differences that set economists and statisticians at odds. I will show that structural parameters indeed correspond to conditional expectations and that zero coefficients correspond to genuine independencies, albeit in nonstandard probability spaces. The interpretation can be formulated in two equivalent forms – one in terms of a process model, the other in terms of an intervention model.

2.1 Process-model interpretation

Note that the distribution corresponding to the structural equations given in (1) is the equilibrium distribution of the following stochastic process:

$$y_1(t+1) = a_1 y_2(t) + a_2 x_1(t) + u_1(t+1) \quad (2a)$$
$$y_2(t+1) = a_3 y_1(t) + a_4 x_2(t) + u_2(t+1) \quad (2b)$$

where $u_1(t)$, $u_2(t)$ are (possibly correlated) stationary disturbances, uncorrelated with $X_1(t')$, $X_2(t')$, $Y_1(t'')$, and $Y_2(t'')$ for all t' and all $t'' < t$.

Based on this observation, we posit that writing down structural equations is intended to designate the dependent variables (usually written on the left-hand side of the equations) as occurring a short but finite time interval after the independent variables (usually written on the right-hand side of the equations). In other words, although the equations are written in a static (or so-called *simultaneous*) mode, the causal content of the equations implies a finite time lag between causes and effects, and it is the structure of this causal content that the analyst summarizes when writing down a set of simultaneous structural equations.

An immediate consequence of this formulation is that the coefficients $a_1, ..., a_4$ can be interpreted meaningfully as conditional expectations. For example, it is not hard to show that

$$E[Y_1(t+1) \mid Y_2(t) = y_2, X_1(t) = x_1] = a_1 y_2 + a_2 x_1 \quad (3a)$$
$$E[Y_2(t+1) \mid Y_1(t) = y_1, X_2(t) = x_2] = a_3 y_1 + a_4 x_2 \quad (3b)$$

We also get a better understanding of what is meant by excluding the X_2 term from (1a) and the X_1 term from (1b). The exclusion reflects a genuine, albeit time indexed, assertion of conditional independence:

$$Y_1(t+1) \perp\!\!\!\perp X_2(t) \mid \{Y_2(t), X_1(t)\} \quad (4)$$

Thus, although $Y_1 \perp\!\!\!\perp X_2 \mid \{Y_2, X_1\}$ does not hold in the equilibrium distribution, a temporal version of this independence assertion holds in the distribution of the process leading to that equilibrium.

It is not unreasonable, therefore, to suppose that economists in writing down (1a) are focusing attention on the dynamic process by which the quantities are generated and that what gives structural equations their "natural meaning" are conditional expectation and conditional independence relations of the sort described in (3) and (4). Thus, while it is true that the process distribution is, generally speaking, unobservable (only the equilibrium distribution can be estimated from the data), we now have a formal interpretation of what Goldberger meant by "the exclusion of x_2 from the demand equation reflects the economist's understanding of household behavior." Eq. (4) makes this understanding explicit and permits us to analyze such understandings using the calculus of conditional independence [3, 19]. Note, for example, that one can make arbitrary changes in the coefficients of (2b), as well as introduce a correlation between u_1 and u_2, without changing the conditional independence asserted in (4).

The process-model account explains why economists who have substantive knowledge of the data-generating process would consider structural equations more natural than regression analysis. These economists find structural equation models more stable and modular than regression analysis models because judgments about one equation can be made independently of judgments about another, which is not the case in regression analysis. As we add more equations, the interpretation of $a_1 y_2 + a_2 x_1$ (in terms of conditional expectation) remains stable relative to the temporal distribution, as in (3a), although the interpretation may change relative the equilibrium distribution. Thus, for process-minded economists, Wermuth and Cox's warning that "the interpretation of equation parameters, be they present or missing, has to be derived from scratch for each model considered" simply does not apply.

2.2 Intervention-model interpretation

This interpretation invokes a nonstandard notion of *conditioning*, where by "conditioning on $X = x$" we mean "holding X fixed (at x) by external intervention." This notion of conditioning, known in the philosophical literature as "imaging" [13], is not common in statistics, because interventions normally change the character of the population under study and, hence, cannot be modeled by Bayes rule

$$P(y \mid x) = \frac{P(y, x)}{P(x)} \tag{5}$$

which applies only when x stands for outcomes of passive observations. The distinction is clear if we consider the difference between "I have observed the barometer reading to be x" and "I intervened and set the barometer reading to x"; we would not consider these two reports equally informative about an incoming storm (y).

Recent advances in graphical and causal modeling [17, 18] show that conditioning on external interventions can be given a precise formal definition that is similar to the definition for conditioning on passive observations. In [17] I used the notation $P(y \mid set(X = x))$ to characterize the distribution resulting from externally fixing $X = x$ and showed that $P(y \mid set(X = x))$ can be readily obtained from the equilibrium distribution whenever a causal structure is given, for example, in a form of a causal network. The analysis rests on

Simon's [20] mechanism-based account of causation and exploits the fact that the intervening mechanism itself can be treated as a variable in a higher dimensional probability space.

While the formalization of $P(y \mid set(X = x))$ in [17, 18] was given in terms of graph operations, it can easily be translated to structural equations models (which, in essence, define directed graphs). In fact, given a set of structural equations, $P(y \mid set(X = x))$ is related to $P(y \mid X = x)$ by a very simple formula:

$$P(y \mid set(X = x)) = P^*_X(y \mid x) \tag{6}$$

where P^*_X is the distribution associated with a diminished set of equations, one obtained by removing the (unique) equation in which X is the dependent variable. The generalization to multiple interventions and observations is straightforward.

The reasoning behind (6) is straightforward: The very essence of the external intervention $set(X = x)$ is to replace the causal mechanism currently governing X with the intervening mechanism whose effect is to set $X = x$. This amounts to removing the equation having X on the left-hand side from the system, turning X into an exogenous variable, and substituting x for X in the remaining equations. Clearly, if X is an exogenous variable, no equation is removed and $P(y \mid set(x)) = P(y \mid x)$. In (1), for example, the calculation of $P(Y_1 = y_1 \mid set(Y_2 = y_2), set(X_1 = x_1))$ amounts to removing (1b) from the system and then computing the value $P(Y_1 = y_1 \mid Y_2 = y_2, X_1 = x_1)$ as dictated by (1a) alone. This procedure yields

$$E(Y_1 \mid set(Y_2 = y_2),\ set(X_1 = x_1)) = a_1 y_2 + a_2 x_1 \tag{7}$$

and

$$Y_1 \perp\!\!\!\perp X_2 \mid \{set(Y_2), set(X_1)\} \tag{8}$$

as expected (compare with (3) and (4)).

Statisticians are bemused by the sensitivity of economists to any changes in the format of the equations of a model; to economists, moving a term between the two sides of an equation seems to change the meaning of the entire system. Is this syntactic sensitivity justified? If so, what (extra) message is communicated through the syntax of the equations that is not conveyed by their solution? One such message is identified quite clearly in (6): the syntax serves to identify the mechanism (i.e., the equation) that should be overridden when we intervene externally and set $X = x$. This information would be lost by allowing arbitrary syntactic transformations. In general, when a set of structural equations is transformed into an algebraically equivalent form, the stationary distribution remains the same and, therefore, no predictions concerned with the results of passive observations are altered. Predictions concerned with the results of interventions will be altered, however (unless the equations are recursive).

Thus the intervention-process account sheds new light on the Goldberger-Wermuth dispute (Section 1). Oddly enough, each author is right and each gives a different interpretation to the sentence "if y_2 and x_1 were fixed." Goldberger takes "fixed" to mean "determined by intervention," while Wermuth takes it to mean "determined by observation." "Determining Y_2 by intervention" renders (1b) invalid, since "intervention" is defined as forcing the equality $Y_2 = y_2$ no matter what values are taken by Y_1, X_1, and u_1. In comparison, "determining Y_2 by observation" implies no external interference with the system; hence, (1b) must remain valid, and Wermuth's result prevails. Thus, Goldberger is right in interpreting (1a) as

$$E(Y_1 \mid Y_2 \text{ and } X_1 \text{ are held fixed at } y_2 \text{ and } x_1) = a_1 y_2 + a_2 x_1$$

and Wermuth is right in noting that

$$E(Y_1 \mid Y_2 = y_2, X_1 = x_1) \neq a_1 y_2 + a_2 x_1$$

The two interpretations are not inconsistent, because

$$\begin{aligned} E(Y \mid X = x) &= E(Y \mid X \text{ is observed to be } x) \\ E[Y \mid set(X = x)] &= E(Y \mid X \text{ is held at } x) \end{aligned}$$

and, in general,

$$E(Y \mid X = x) \neq E[Y \mid set(X = x)]$$

3 Remarks

The process-model interpretation is not foreign to the econometric literature. Wold and Jureen [26], for example, provide an in-depth time-series analysis of the demand equation, which involves both lagged and instantaneous causation. Malinvaud [15] and Strotz and Wold [21] have explicitly interpreted causal cycles in terms of recursive models with time lags. However, these temporal process models were presented as a refinement, rather than an explicit definition of structural equations methodology, and so they have not been construed by statisticians as convincing justification for this methodology.

The intervention-model interpretation can be traced to Haavelmo [10], who writes:

> What *is* then the significance of the theoretical equations obtained by omitting the error terms...? To see that, let us consider, not a problem of passive predictions, but a problem of government planning. Assume that the Government decides, through public spending, taxation, etc., to keep income, r_t, at a given level.

Haavelmo then shows that this assumption yields the desired conditional expectation interpretation for the systematic part of the equation and concludes that "this is only natural, because now the Government is, in fact, performing 'experiments' of the type we had in mind when constructing each of the two equations." Clearly, then Haavelmo believes that when an economist writes down structural equations, he/she has in mind controlled experiments.

Freedman [6], one of the most articulate critics of structural equations, has also embraced the intervention interpretation: "My view, stated in detail earlier, is that a path model represents the analysis of observational data as if it were the result of an experiment." However, by his own admission, Freedman's interpretation remains vague and informal: "At points such as this, it would be helpful to know more about the structure of such hypothetical experiments: What is to be held constant, and what manipulated?" Indeed, Freedman fails to recognize, for example, that an intervention should always be modeled by removing the equations in which the manipulated variables are dependent [6]. This failure is not so much the fault of Freedman as it is an indictment of the econometric literature for not explicating formally either what sort of experiment is modeled by structural equations or how structural equations are tied to the calculus of intervention.

Thus, although both the intervention and temporal accounts have previously been considered and accepted by economists, the controversies and misunderstandings (such as those described above) that keep flaring up in the contemporary literature indicate that these accounts have not been formulated with a clarity and precision sufficient to establish the legitimacy of structural equations methodology. Regardless of the methods we use for testing, identifying, or inferring structural equations models, we must first understand what the structural equations model stands for, what claims it communicates, and how it is being used. I hope this note will help bring about such understanding.

References

[1] Bentler, P.N., "Multivariate analysis with latent variables," *Review of Psychology*, 31, 419-456, 1980.

[2] Blalock, H.M., Jr., *Causal Inference in Nonexperimental Research*, University of North Carolina Press, Chapel Hill, 1964.

[3] Dawid, A.P., "Conditional independence in statistical theory," *Journal of the Royal Statistical Society, Series A*, 41, 1-31, 1979.

[4] DeLeeuw, J., "Review of four books on causal analysis," *Psychometrika*, 50, 371-381, 1983.

[5] Duncan, O.D., *Introduction to Structural Equations Models*, Academic Press, New York, 1975.

[6] Freedman, D., "As others see us: A case study in path analysis" (with discussion), *Journal of Educational Statistics*, 12, 101-223, 1987.

[7] Goldberger, Arthur S., *Structural Equation Models in the Social Sciences*, Seminar Press, New York, 1973.

[8] Goldberger, Arthur S., "Models of substance; comment on N. Wermuth, 'On block-recursive linear regression equations'," *Brazilian Journal of Probability and Statistics*, 6, 1-56, 1992.

[9] Goldberger, A.S., *Econometric Theory*, John Wiley, New York, 1964.

[10] Haavelmo, T., "The statistical implications of a system of simultaneous equations," *Econometrica*, 11, 1-12, 1943.

[11] Jöreskog, K.G., and Sörbom, D., *LISREL IV: Analysis of Linear Structural Relationships by Maximum Likelihood*, International Educational Services, Chicago, 1978.

[12] Keynes, J.M., "Professor Tinbergen's method," *Economic Journal*, 49, 560, 1939.

[13] Lewis, D.K., "Probability of Conditionals and Conditional Probabilities", *The Philosophical Review*, 85, 297-315, 1975.

[14] Ling, R., Review of *Correlation and Causation*, by D. Kenny, *Journal of American Statistical Association*," 489-491, 1983.

[15] Malinvaud, E., *Statistical Methods in Econometrics*, trans. A. Silvey, North-Holland, Amsterdam, 1978.

[16] Niles, H.E., "Correlation, causation, and Wright theory of 'path coefficients'," *Genetics*, 7, 258-273, 1922.

[17] Pearl, J., "Aspects of graphical models connected with causality," In *Proceedings of the 49th Session of the International Statistical Institute*, Tome IV, Book 1, Florence, Italy, 391-401, 1993.

[18] Pearl, J., "Comment: Graphical models, causality, and intervention," *Statistical Science*, 8:3, 266-269, 1993.

[19] Pearl, J., *Probabilistic Reasoning in Intelligent Systems*, San Mateo, CA, Morgan Kaufmann, 1988.

[20] Simon, H.A., *Models of Discovery: and Other Topics in the Methods of Science*, D. Reidel, Dordrecht, Holland, 1977.

[21] Strotz, R.H., and Wold, H.O.A., "Recursive versus nonrecursive systems: An attempt at synthesis," in H.M. Blalock (Ed.), *Causal Models in the Social Sciences*, Aldine Atherton, Chicago, 1971.

[22] Tinbergen, J., *An Econometric Approach to Business Cycle Problems*, Hermann, Paris, 1937.

[23] Wermuth, N., "On block-recursive regression equations" (with discussion), *Brazilian Journal of Probability and Statistics*, 6, 1-56, 1992.

[24] Wermuth, N., and Cox, D., "Linear dependencies represented by chain graphs," *Statistical Science*, 8:3, 204-218, 1993.

[25] Whittaker, J., *Graphical Models in Applied Multivariate Statistics*, John Wiley, Chichester, England, 1990.

[26] Wold, H., and Jureen, L., "The ends and means of demand analysis," in *Demand Analysis*, pages 1-20, John Wiley, New York, 1953.

[27] Wright, S., "Correlation and causation," *Journal of Agricultural Research,* 20, 557-585, 1921.