

# ACTIONS, CONSEQUENCES, AND CAUSAL RELATIONS<sup>1</sup>

Guy H. Orcutt

WE regard economic phenomena as resulting from an interaction of human actions within a field of nonhuman environmental conditions and restraints. Of the various phases of economic phenomena that may interest social scientists there is one which is of particular importance to policy-makers whether they are in business or in government. One of their special needs is for knowledge concerning the consequences or impacts of the various actions which they are able to take and which they consider taking. This paper deals with methods that more adequately attempt to meet this need.

We assume that the policy-makers know what they are seeking to achieve. We also assume that they have a number of actions at their disposal and that they wish to know which of these will achieve the desired objectives in as satisfactory a manner as possible. We take it for granted that it will usually be necessary to observe discrepancies between the desired situation and the actual situation in order to guide whatever actions are taken, but we will not concern ourselves here with problems of this sort. Rather we will limit our attention to certain aspects of the problems involved in discovering and specifying the consequences of actions.

## *Causal relations defined*

This paper deals primarily with the inference of causal relations suitable for specifying the consequences expected from actions. However,

<sup>1</sup>The author has received considerable assistance from discussions with R. Cameron, W. Fellner, G. Haberler, E. Kuh, W. Leontief, P. Samuelson, H. Simon, and R. Solow.

Recent works which may be of interest to the reader in this field include the following: T. C. Koopmans, "When Is a System Complete for Statistical Purposes?" in ch. 17 of *Statistical Inference in Dynamic Economic Models*, ed. T. C. Koopmans (Cowles Commission Monograph No. 10, New York, 1950). Guy H. Orcutt, "The Inference of Causation," read before the September 1950 meetings of the Econometric Society, abstract in *Econometrica*, XIX (1951) 60-61. Arthur W. Burks, "The Logic of Causal Propositions," *Mind*, LX (July 1951), 363-82. Herbert A. Simon, "On the Definition of the Causal Relation," mimeographed (1952); and Herbert A. Simon, "Causal Ordering and Identifiability," ch. III of *Econometric Methods*, ed. Wm. C. Hood and T. C. Koopmans (Cowles Commission Monograph No. 14, John Wiley and Sons, New York, forthcoming).

let us first discuss the meaning that we attach to the term causal relation and why it is important to look for causal relations.

By causal relations we mean asymmetrical or unidirectional relations such as, "If  $A$ , then  $B$ ." For example we regard the statement, "If it rains, there will be puddles," as a causal relation. We do not take this statement to say anything at all about whether or not rain will be observed if puddles are observed. Nor do we necessarily take such a statement to say anything about whether or not rain will be observed if puddles are not observed. Since, by definition, causal relations are taken to be unidirectional, it is clear that they are more restricted in nature than relations which are understood to be of a symmetrical or nondirectional nature. To express one nondirectional relation in terms of causal relations requires at least two causal relations. For example, suppose we wish to express a nondirectional relation,  $\Delta X = b \Delta Y$ , in terms of causal relations. To do this we would have to say, "If  $Y$  changes by  $\Delta Y$ , then  $X$  will change by  $b \Delta Y$ ," and "If  $X$  changes by  $\Delta X$ , then  $Y$  will change by  $\frac{1}{b} \Delta X$ ." Since

either of these two causal relations might be regarded as true without regarding the other as true, it is apparent that a single causal relation between variables is a more restricted type of statement, in the sense that less is claimed than is claimed by a nondirectional relation between the same variables.

There is one other, not entirely independent, difference between causal and nondirectional relations which we would like to include as part of our description of causal relations. Nondirectional relations are usually thought of as being either essentially complete and exact, or of holding with well defined limits of either an absolute or statistical sort. Causal relations may sometimes be set up as unidirectional relations which are complete and exact within well defined limits, but they need not be set up this way. Thus, when we say that  $A$  is a cause of  $B$ , we often mean that if  $A$  varies, then  $B$  will be

different in a specified way from what it *would have been* if *A* had not varied. We do not, however, exclude the possibility that *B* would have changed by an unknown amount even in the absence of any variation of *A*. One example of such a causal statement would be, "If this light is turned on, the light intensity at a particular place in this room will be (within certain narrow limits) a specified amount more than it otherwise *would have been*." Or again, "If this man's income is increased, his consumption will be larger than it otherwise would have been."

#### *Why look for causal relations?*

We are interested in predicting the consequences of human actions, and to do so we must find relationships between economic phenomena and actions or between some economic variables and other economic variables that we know how to control and in terms of which actions might be specified. We seek to establish such relations on the basis of previous experience and then, taking account of all factors found to be of importance, we use these relations for predicting the future relation between consequences and actions. But why emphasize the importance of looking for unidirectional, or causal, relations rather than simply nondirectional relations?

There are three essential reasons why, in seeking to discover the impact of actions, we should look for causal relations. First, as emphasized in a preceding paper,<sup>2</sup> the policy implications of any relation depend critically upon whether it holds in one or more directions. Thus given a relation between consumption and income it is important to know whether it should be expected to hold when actions controlling income are used, when actions controlling consumption are used, or in both of these cases. Secondly, it seems clear that methods which are not designed to recognize the directional nature of relations will often lead to acceptance of a relation as nondirectional when, on the basis of available data, only a more restricted causal relation is justified. And finally, it also seems clear that if we do not use techniques adapted to finding causal, as contrasted with nondirectional, relations, we may fail to

find relations which actually exist and which could be found on the basis of available data.

#### *Idealized experiments by a single experimenter*

It has sometimes been said that all we ever observe between real variables are associations. If this is true, what kind of experience ever suggests the use of unidirectional relations in reference to real phenomena? To see one way in which we are led to set up unidirectional rather than nondirectional relationships on the basis of sets of observed associations, consider the following.

We abstract from some of the difficult problems of induction, and so focus attention on what is crucial for our present purposes, by considering the case of an experimenter dealing with variables that he alone acts upon. This being the case and given sufficient time and energy, he might try every action and every combination of actions that he knows. Let us suppose he treats each action and each combination of actions as separate actions, and that he gives a number to each one. Each time that he tries some particular action he would observe all the objective variables that interested him, and then enter the observed results of the action in a large table opposite the number of the action. A section of his table might resemble Table I. We have merely put crosses under

TABLE I

Action	Objective Variables						
	Z <sub>1</sub>	Z <sub>2</sub>	Z <sub>3</sub>	Z <sub>4</sub>	Z <sub>5</sub>	—	—
1	x	x		x			
2		x	x				
3		x			x		
4	x	x	x				
5	x				x		
—							
—							

each variable for which he observes a response corresponding to each action. Actually, of course, we would expect our experimenter not only to observe what variables responded, but also the way in which they responded at each intensity of the action. This added detail would evidently introduce the possibility of much more complicated causal structures, but for the purpose of illustrating our point the crosses will be sufficient.

<sup>2</sup> Guy H. Orcutt, "Toward Partial Redirection of Economics," this REVIEW, xxxiv (August 1952).

Our experimenter now has relations between each of his actions and some objective variables. For his own convenience he might well wish to systematize or fit into some theoretical structure his cook-book type of knowledge. Considerations of the type that would arise in doing this might indeed lead him to construct causal relations or unidirectional relations. However, let us leave this possible approach to causal relations and consider the problem of transferability of knowledge.

We may suppose that there are other individuals who could carry out some or all of the actions of the experimenter. However, while they have learned to associate their actions with certain subjective and objective consequences, they have not learned the relation between their actions and many objective variables of interest. Now it would be nice, of course, if our experimenter could specify an action independent of its objective results. Then having specified actions 1, 2, etc., he could say action 1 affects, in such and such a way, objective variable  $z_1$ ,  $z_2$ , and  $z_4$ . Action 2 affects, in such and such a way, objective variables  $z_2$  and  $z_3$ . And so on down the list. However, the shortcomings of this kind of procedure are obvious. The experimenter may possibly be able to recognize his own actions independently of objective variables that are observable by others, but no one else can. The experimenter could tell other individuals that there is an action which affects  $z_1$ ,  $z_2$ , and  $z_4$ . There is also another action which affects  $z_2$  and  $z_3$ . And so on. This, however, would be of rather limited usefulness. To convey more useful information one must specify actions in terms of only part of the objective variables which they act upon. If this can be done, then it would be possible to say to someone that if you know an action that will control some particular variables, use it and you will discover that it will also control such and such other variables. For example, look at Table 1 and consider just the first four actions. Within this field of possibilities one would be right in saying, use an action which controls  $z_1$  and you will also have an action which controls  $z_2$ . But notice that the converse statement is not correct. As a matter of fact, it appears that there are many actions which control  $z_2$  but which do not in turn control  $z_1$ . Looked at from this

point of view, we see that the statement that  $z_1$  is in a causal chain leading up to  $z_2$ , or that  $z_1$  is a cause of  $z_2$ , is just a convenient way of saying that if you pick an action which controls  $z_1$ , you will also have an action which controls  $z_2$ . But, as is apparent from the above, the converse may or may not be true. All we ever observe are associations. But the associations include associations between actions and objective variables. The need for causal relations arises here because we can only specify an action to someone else in terms of part of what the action does to objective variables which they can observe. To be satisfactory, causal relations must in effect give correct instructions as to what will happen when various objective variables or sets of objective variables are controlled. The asymmetry of such instructions arises because, as in the above example, to say "Pick an action which controls  $z_1$ " does not limit the choice to the same set of actions as does the instruction "Pick an action which controls  $z_2$ ."

It may however happen that someone succeeds in controlling  $z_1$  but is disappointed in his expectation that he will thereby also control  $z_2$ . What the accepted associations showed was that, as far as the body of experience taken into consideration is concerned, all methods of varying  $z_1$  are associated with varying  $z_2$ , but that some method of varying  $z_2$  exists which does not result in variation in  $z_1$ . They could not, of course, show that we were correct in acting as though they were a complete set of associations. The discovery that  $z_1$  could be controlled without  $z_2$ , as under action 5 of Table 1, would lead to an enlargement of our field of accepted associations and would call for a revision of our causal relations. It might also spur a search for one or more additional objective variables associated with the techniques that work, but not with those that do not work, in controlling  $z_2$ . In this particular case we would be correct if we said, "Pick an action which controls  $z_1$  and  $z_4$  and you will have an action which controls  $z_2$ ." Our hope is that, eventually, as our knowledge of associations increases, we will be able to construct a causal system which will yield correct predictions no matter who makes use of it, or what techniques are used to control the variable or combination of variables in terms of which actions are specified.

*Nonidealized experimentation*

In real situations the experimenter can recognize and distinguish his own actions with sufficient accuracy, and he can still carry them out in isolation or combination. He can also observe the behavior of objective variables to see what happens. The big difference is the presence of many other sources of disturbance besides himself. Now the experimenter must develop strategies aimed at sorting out the effect of his own action from that of other disturbances. He designs experiments which are carried out in such a way that some variables are held constant. He carries out his experiments in different places at different times of the day, etc. If there are variables he would like to hold constant but cannot, he may carry out his experiment at different observed values of these variables, and thus still separate out the effect of his action. In cases in which he can neither control the other disturbances nor observe them he can then repeat his action many times and in an irregular time sequence so as still to frequently succeed in isolating the impact of his own action from that of the noncontrolled and nonobserved disturbances. The experimenter never does succeed in varying one variable while observing another and at the same time holding everything else constant. Nevertheless, he has been very successful at discovering useful causal relations.

*Personal observation*

But what about the economist? We are seldom able to avail ourselves of experimental procedures. Or perhaps I should say that we seldom do avail ourselves of experimental procedures, since clearly much more is possible in this direction than has been done. To the extent that we use engineering data as Leontief and his associates are doing, we are availing ourselves of experimental results, and the associations so obtained may assist greatly in the construction of useful causal structures. However, what about the use of nonexperimental data, upon which we must put our main reliance?

In the first place, each of us does have a lot of experience which is identical in nature with

that gained by experiment. Thus, we all are continually taking actions, and we all do a certain amount of observing. The only differences are that we do not usually select a sequence of actions and the time and place of performing them with a view to studying and isolating their impact, and usually we do not take great pains in observing their impact. Nevertheless, it is clear that for the careful observer this source of experience may be and has been useful in constructing causal relations in precisely the same way as experimental evidence is useful. In fact, it would be reasonable to classify it under experimental data.

*Observational data — Actions indirectly observed*

Up to this point it has been reasonably clear sailing, because the observer, being the key actor, has been well situated from the standpoint of associating actions with impacts on objective variables. Difficulties begin to mount as soon as the observer tries to learn from the actions of others. This is, of course, usually our position when we are attempting to construct causal relations on the basis of published data. We still can search for and sometimes discover associations or relations between objective variables, but how do we tie these associations to actions? To do this we must build on the base provided by what we have learned from our observations of the impact of our own actions. We have learned to associate certain aspects of objective experience with each action. We now use these elements of objective experience as pointers to the actions we associated with them. If others in following our instructions do something which produces the objective manifestations we associate with an action, we go ahead on the assumption that their action is identical or nearly identical with that action which we associate with these objective manifestations. Sometimes later events prove that we were wrong, and then a further modification of ideas about the specification of actions in terms of objective variables must take place. This process continues until for most practical purposes we correctly learn to recognize actions, not only in terms of the way they "feel" to us when carrying them out, but

also in terms of objective manifestations that can be observed when either we or other individuals carry them out. From a practical point of view the recognition is correct as long as actions considered the same actually do have sufficiently similar impacts on all variables of interest.

In so far as we do in fact succeed in specifying and recognizing actions on the basis of objective phenomena, the way is open for the extension of our causal relations by inferences based upon purely observational data. We again are in a position of being able to recognize when an action has been taken. We again may find associations or relations between objective variables that we also associate with particular actions.

But before we proceed too rapidly it is perhaps worth while considering under what circumstances we feel fairly safe in saying that we can in fact recognize an action by its objective manifestations. Can we, for instance, look at aggregate time-series data and recognize actions? The number of actions which simultaneously determine highly aggregative variables are so numerous that we are almost in the position of observers in fields in which actions are not taken and we can only observe. In such fields causal relations may sometimes be appropriate but not in relation to the specification of actions. Economic phenomena, however, are clearly modified by human actions, and causal relations may be appropriate for the reasons developed up to this point. However, to discover them we must get down to a degree of "dis-aggregation" where we have a reasonable chance of finding associations between objective variables which can also be related to actions. To treat our data as great aggregates is to lose nearly all connection between variables and actions except in the case of a few "experiments" on a grand scale by governments or other large units. These "experiments" may indeed have an impact sufficiently large to permit partial separation of their impacts from the impacts of other noncontrolled events. However, these "experiments" on a grand scale, while of great importance, usually are not repeated often enough to make possible much in the way of inference unless a great deal has

already been learned on the basis of other kinds of experience. Nor is it only that "experiments" by governments or other large units are usually either too small and diffused or too few. They may in fact be systematically carried out in such a manner that they never give us the experiments we need most. For example, if exchange rates were systematically adjusted so as always to achieve a balance, then changes in exchange rates would always coincide with and counter some other disturbance. This being the case, it would be very difficult to separate out the effect of changing the exchange rate unless we knew how to allow for the effect of the disturbing factors.

Our position then is that, while making all possible use of aggregate time series, we must nevertheless place our main hopes for the determination of causal relations upon less aggregative data. This would seem to leave only cross-section data and less aggregative time-series data. The crucial importance of both of these types of data is being increasingly recognized and every effort should be made to exploit them. Nevertheless, from the viewpoint of putting variables into correct positions in causal relations, I feel that either less aggregative time series or a combination of these two types of data will prove to be most useful. This I think is true because we are probably in a much better position to recognize an action on the basis of the behavior of variables over time than on the basis of purely cross-section evidence. Our own actions usually extend over time, and the impacts are also spread out over time. Furthermore, it seems likely that the impact of an action will appear not so much in the form of a configuration at a point in time of particular values of variables but rather in the form of changes in the value of variables over time.

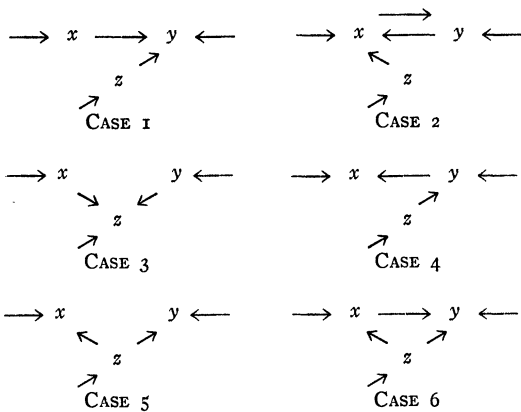
#### *Observational data — Actions not observed*

We have emphasized that, in determining causal relations aimed at relating actions to their consequences, it is important that the investigator be able either directly or indirectly to observe the action that is connected with the association of objective variables under study. But, while stressing the importance of this, it is also worth noting that some limited infer-

ences about causal structure can be made upon the basis of data even where it is not possible directly or indirectly to observe actions independently of some or all of the variables being related.

One commonly used way in which we infer something about causal relations in cases in which actions are not observed involves taking cognizance of time relations. If we observe an association between two variables then we may be willing to regard the variable which lags in time as in a causal chain leading from the variable which precedes in time, but not *vice versa*. In doing this we are making use of one of the most general inferences that has been made from all the experimentation that has been done: namely, that actions taken in the present do not appear to modify the past. Since in all cases in which we do control variables we have apparently never been able to alter what has already happened, it seems reasonable to construct our causal relations on the assumption that this will also be true of actions taken in the future.

Next, consider the case of three economic time series,  $x$ ,  $y$ , and  $z$ . Suppose, on the basis of the available historical evidence, it is decided that  $x$ ,  $y$ , and  $z$  are linearly related. This taken by itself would be compatible with many causal schemes including the following ones.



In Case 1 the notation is interpreted as follows. There are actions which affect  $x$ , actions which affect  $z$ , and actions which affect  $y$ . The actions which affect  $x$  also affect  $y$  but not  $z$ . The actions which affect  $z$  also affect  $y$  but not  $x$ . The remaining actions which affect  $y$  do not

affect  $x$  or  $z$ . However, notice how the acceptable causal structures are greatly reduced, if, on the basis of the historical evidence, it is decided that not only are  $x$ ,  $y$ , and  $z$  linearly related, but that the behavior of  $x$  has been independent of the behavior of  $z$ . Now, out of those causal structures shown only the first one is still compatible with the evidence. All the others imply either that an action which affected  $z$  would also affect  $x$ , or that an action which affected  $x$  would also affect  $z$ . However, we are not so far advanced nor on such secure ground as if we could actually identify the actions in question. We may not know any actions for controlling any of these variables. Or if we do they may not be at all analogous in their impacts to those actions or circumstances which were of importance in generating the data. Nevertheless, it is clear that in cases in which the choice can be narrowed down on the basis of evidence of the type discussed earlier, the possibility does exist of using this type of evidence in making inferences about causal relations aimed at relating actions to their consequences.

Another type of three-or-more-observed-variable case in which something might also be done is one in which we accept two or more variables as being in the same causal chains leading up to another variable but do not know in which order to place them. Thus, consider the following:

1.  $\rightarrow x \rightarrow y \rightarrow z$   
           $\nearrow$            $\nearrow$
2.  $\rightarrow y \rightarrow x \rightarrow z$   
           $\nearrow$            $\nearrow$

The symbolism in the first case is to be read as follows. There exist separate actions which affect  $x$ ,  $y$ , and  $z$  respectively. Those actions which affect  $x$  also affect  $y$ , but not all actions affecting  $y$  affect  $x$ . Furthermore, those actions which affect  $y$  affect  $z$ , but not all actions affecting  $z$  affect  $y$ . Now, supposing that historical data are available which give a high multiple correlation of  $z$  on  $x$  and  $y$ ; a high partial correlation of  $z$  on  $y$  with  $x$  as the third variable; and essentially a zero partial correlation of  $z$  on  $x$  with  $y$  as the third variable. In this case

the evidence would be quite consistent with the first causal structure but inconsistent with the second.

In the preceding cases we were able to say something about whether  $x$  is in a causal chain leading up to  $y$  or *vice versa*, but in order to do so we required that a third variable  $z$  also be observed. What about the possibilities if only two variables,  $x$  and  $y$ , are observed and we again are unable to observe independently the actions which are linked to the variations of  $x$  and  $y$ ? The possibilities of making inferences about causal structure in this case are more limited than in the preceding case, but even here something is possible if the variations of  $x$  and  $y$  are of the right sort. Thus, for example, suppose that the scatter diagram of the available pairs of  $x$ 's and  $y$ 's looks like in the scatter diagram of Chart I.

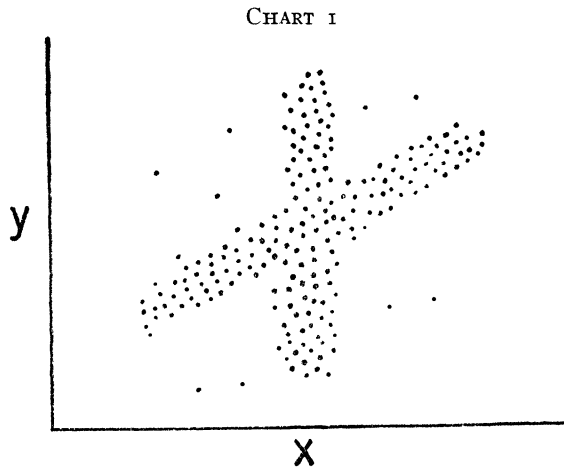


CHART I

This evidence would be consistent with the causal structure  $\rightarrow x \rightarrow y$ , where actions which affect  $x$  also affect  $y$  but other actions exist which affect  $y$  but not  $x$ . However, the scatter diagram would not be compatible with many other types of causal structure. For example, consider it in relation to the following causal structures.

1.  $\rightarrow x \leftarrow y \leftarrow$
2.  $\rightarrow x \rightleftharpoons y \leftarrow$
3.  $\rightarrow x \begin{matrix} \swarrow \\ \downarrow \\ \searrow \end{matrix} y \leftarrow$

However, while inferences of the above type may on occasion be possible, it is perhaps worth noting that, insofar as certain common assumptions about normality are true, bivariate distributions of the type shown in Chart I would never arise. Thus, suppose that,

$$y = a + bx + \epsilon$$

where  $x$  and  $\epsilon$  are distributed independently and  $\epsilon$  is not observed. Then, if  $x$  and  $\epsilon$  are both normally distributed random variables, it is evident that the joint distribution of  $x$  and  $y$  will be a bivariate normal one. In this case, even though the correct causal structure really is

$$\rightarrow x \rightarrow y \leftarrow$$

the bivariate distribution will be equally compatible with

$$\rightarrow x \leftarrow y \leftarrow$$

In this particular case, as long as actions controlling either  $x$  or  $y$  are not taken, it perhaps does not matter which is the correct causal structure. But if we wish to predict what will happen to  $y$  if actions controlling  $x$  are taken, then we do need to know the correct causal structure.

One type of data which might yield evidence of the sort indicated in Chart 1 would be time series exhibiting well defined, well spaced, and irregularly occurring pips. Common pips on the two series would support the belief in some connection between them. Extra pips on one but not the other would be compatible with causality running from the variable without extra pips to the variable with extra pips but not *vice versa*. Extra pips on both series would be compatible with a common cause but no direct causal link, and so on.

*Koopmans' views on causal structure*

In the symposium in the previous issue of this REVIEW it became evident that Koopmans and I shared a somewhat similar view of the meaning of causal relations.<sup>3</sup> Nevertheless, Koopmans'

<sup>3</sup> Orcutt, "Toward Partial Redirection of Econometrics," and "Comment" by Koopmans, op. cit. See also Koopmans, "When Is a System Complete for Statistical Purposes?" in *Statistical Inference in Dynamic Economic Models*, op. cit.

position with regard to the possibility of inferring causal structures seemed to be much more pessimistic than my own. If I am correctly interpreting Koopmans, it will become apparent that his pessimism stems from limiting his attention to by far the most unfavorable case that we have discussed. The model Koopmans works with is as follows. Assume a complete system of all structural equations between all variables, economic or noneconomic, that enter directly or indirectly into the explanation of economic variables. Suppose that there are  $N$  of these variables and that they all enter without time lags. Further assume that the equations and the variables can be separated into two sets with numbering  $n = 1, \dots, G$  and  $n = G + 1, \dots, N$  for the first and second sets respectively, such that the first set of endogenous variables does not occur in the second set of equations. Finally, assume that the random terms, of which there is one in each equation, are such that those entering the first set of equations are distributed independently of those entering the second set. Thus, we are to assume a system such as the following:

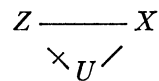
	Endogenous variables	Exogenous variables	Random terms
	$f_1(x_1, \dots, x_G;$	$x_{G+1}, \dots, x_N)$	$= u_1$
	$\cdot$		$\cdot$
Set I	$\cdot$		$\cdot$
	$\cdot$		$\cdot$
	$f_G(x_1, \dots, x_G;$	$x_{G+1}, \dots, x_N)$	$= u_G$
		$f_{G+1}(x_{G+1}, \dots, x_N)$	$= u_{G+1}$
		$\cdot$	$\cdot$
Set II		$\cdot$	$\cdot$
		$\cdot$	$\cdot$
		$f_N(x_{G+1}, \dots, x_N)$	$= u_N$

in which the  $u_1$  through  $u_G$  are assumed to be distributed independently of the  $u_{G+1}$  through  $u_N$ . The feature which characterizes the exogenous variables is that while their variations may affect the endogenous variables, variations of the endogenous variables for other reasons do not react back on the exogenous variables. When I first looked at Koopmans' model I could not help wondering why, if Koopmans wanted to assume a complete system of structural equations, he persisted in retaining the nonexplained random terms. But after considerable reflection it became clear that his specification of the independence of the two sets of random terms, far

from being a superficial element, really lay at the heart of his development, was analogous to my own independent development, and furnished the only basis of testing any choice of exogenous variables within the framework of his model. Without the  $u$ 's it is clear that experience could not in general offer any test of which variables are to be found in set II and not in set I. For, given a set of  $N$  exact equations with  $N$  variables, it is usually possible to solve for any subset of the variables in terms of the remaining ones and thus obtain a subset of equations from which certain variables have been eliminated. Since all linear transformation of the original set of equations would fit perfectly, there would never be any basis of choice as to which variables would end up in set II. We may simplify our discussion of Koopmans' model by looking at it as specifying relations between only four vector variables. Thus, let:

- $X$  represent  $x_1, \dots, x_G$
- $Z$  represent  $x_{G+1}, \dots, x_N$
- $U$  represent  $u_1, \dots, u_G$
- and  $W$  represent  $u_{G+1}, \dots, u_N$ .

We can drop consideration of  $W$  as soon as we can recognize that since  $Z$  is entirely specified by specifying  $W$  and since  $W$  is assumed to be distributed independently of  $U$ , that it follows that  $Z$  is also assumed to be distributed independently of  $U$ . Thus, we arrive at the following system



with variations in  $Z$  related to variations in  $X$ , variations in  $U$  related to variations in  $X$ , but no relations between the variations of  $Z$  and  $U$ . Now, given that  $Z$  and  $U$  vary in complete independence, and assuming that the associated variation of  $Z$  and  $X$  is due to one action or class of actions while the associated variation of  $U$  and  $X$  is due to another action or class of actions, we would be able to make the following statements. On the assumption that our set of associations is sufficiently complete, if you pick an action which controls  $Z$  you will have one which controls  $X$ . It is also true that if you pick an action which controls  $U$  you will have one which controls  $X$ . However, to say "pick an



action which controls  $X''$  does not specify the action closely enough to say whether you will have one which also controls  $Z$  or  $U$ . Koopmans is thus seen to be dealing in terms of what, as far as causal structure is concerned, is our case in which only two variables are observed and in which actions that may be related to their movements are not independently observed. As we

and others have pointed out, inference of causal relations from the data in this case is impossible under certain common assumptions of normality and is more or less tenuous and of limited scope in any case. Under these highly restrictive conditions pessimism is certainly justified, but there is no reason why we should limit ourselves to these restrictive conditions.