V

In summary we believe that insofar as econometricians wish to be of more service to policy-makers, the following redirection of effort is called for:

1. More attention needs to be directed at determining the impact of those variables which might actually be used for control purposes.

2. More study should be given to the continuity properties of economic time series. This type of study is needed as a basis for specifying what kind of lags can be tolerated in the impact of instruments of control or policy actions.

3. A vigorous effort should be directed at discovery of exogenous variables and of as complete a specification as possible of their impact. Not only do the policy implications of econometric models depend critically on the selection of exogenous variables, but the specification of the estimation processes used also depends upon a proper selection of exogenous variables.

4. Insofar as possible, exogenous variables which move independently of one another should be sorted out but, insofar as the chosen exogenous variables are interrelated, these interrelations must be investigated before an econometric model can be very useful to the policy-maker.

COMMENTS

I

By T. C. Koopmans [1]

DR. ORCUTT has selected a very useful and appropriate model for the discussion of economic policies. His model expresses what A. P. Lerner has called the "economics of the steering wheel." [2] In order to keep the car on the road, it is not necessary to be able to predict, with great accuracy and for a long time ahead, the course the car will follow if the steering wheel is held in a fixed position. Rather, it is sufficient if one can observe departures from the middle of the road, and respond quickly to such observations with changes in the position of the wheel that return the car to the middle of the road. The position of the steering wheel stands for a *controllable* (or *instrumental*) variable,[3] that is, a variable which can be given any value (within a certain range) by the implementation of a policy decision. The proximity to the middle of the road stands for the *objective* variable which is a gauge of the success achieved by the policy.

[1] I am indebted to J. Marschak and H. A. Simon for valuable comments on an earlier draft of this note.

[2] A. P. Lerner, *The Economics of Employment* (New York, 1951), see ch. 1.

[3] The term "controllable exogenous" variable has been used by J. Marschak (*Statistical Inference in Dynamic Economic Models*, ch. 1, section 1.2.3.2); the term "instrumental" variable by J. Tinbergen in what seems to me the same meaning.

I

I find myself in full agreement with Orcutt's proposal that instrumental variables should be incorporated explicitly in econometric models. Nor do I believe that Orcutt and I stand isolated on this point. The same idea can be found in Klein's work,[4] and the use of econometric models to discuss the effects of policies has been developed into a fine art by Tinbergen.[5] It is, I believe, in the best tradition of econometrics to regard econometric model construction as an aid in the discussion of policies. The principal objective is to make

[4] See, for instance, "The Use of Econometric Models as a Guide to Economic Policy," *Econometrica*, April 1947, reprinted as Cowles Commission Paper, New Series, No. 23, in particular p. 112, pp. 138–39.

[5] J. Tinbergen, "Der Einfluss der Kaufkraftregulierung auf den Konjunkturverlauf," *Zeitschrift Für Nationalökonomie*, 1934, pp. 289–319. "Quantitative Fragen der Konjunkturpolitik." *Weltwirtschaftliches Archiv*, November 1935, pp. 366–99. "Über die Sekundärwirkungen zusätzlicher Investitionen." *Weltwirtschaftliches Archiv*, January 1937, pp. 39–57. "Conjunctuurpolitiek en Internationale Verhoudingen," *De Economist*, 1937, pp. 81–107. *An Econometric Approach to Business Cycle Problems*, Paris, 1937 (see ch. III). "On the Theory of Business-Cycle Control," *Econometrica*. January 1938, pp. 22–39. *Business Cycles in the United States of America 1919–1932*, League of Nations, Geneva, 1939. *Business Cycles in the United Kingdom 1870–1914*, Verhandelingen der Koninklijke Nederlandse Akademie van Wetenschappen, AFD. Letterkunde (Amsterdam, 1951), ch. 9.

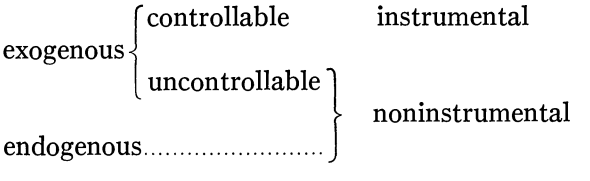
available to the policy-maker any and all knowledge about the effects of alternative policies on economic welfare that can be provided by systematic analysis of observations of a statistical, institutional, or introspective character. In particular, explicit incorporation into econometric models of the dials and levers set by policy decisions is necessary to guide the allocation of research effort to the various aspects and alternative elaborations of the models in question.

Explicit recognition of instrumental variables may also help to dispel the very widespread belief that the main criterion of success in econometric model construction is the ability to predict accurately and for a considerable time ahead the course of economic variables. I cannot help feeling that Dr. Orcutt, himself a distinguished econometrician, is influenced by this belief in phrasing the objectives he imputes to econometricians in the opening sentences of his article, and in speaking of the relative failure of econometrics when the limitations to forecasting possibilities become evident. Econometricians are bound to be frustrated individuals as long as performance in accurately forecasting the future paths of economic variables for an extended period ahead is regarded as the main criterion of success, by themselves or by others. For each type of economic decision only a few important determining factors can be isolated, observed, and assessed as to their influence on the decisions in question. Many other factors affecting economic decisions are so diverse in origin as well as in point of application, and so far outside the range and power of systematic scientific observation, recording, and analysis, that we can only treat their joint effect as a random disturbance, not directly observed. This places very definite though as yet unknown limits on the extent to which scientific analysis can at all forecast economic developments. Without implying that econometric model construction has come anywhere within sight of these limits, awareness of their existence reinforces Orcutt's emphasis on explicit treatment of instrumental variables. The main purpose of econometric model construction is to determine which dials and levers are sufficiently fast and predictable in their effects so that the policy-maker, by adjusting these dials continually in prompt response to the observed path of economic variables, has a more than even chance to steer a more stable, efficient, and growth-enabling course than the economy would take if other or fewer conscious policies were applied. While this may be regarded as the purpose of most of dynamic economic analysis, it is difficult to believe that the insights and precepts of economic theory and experience cannot be made more reliable and more specific, hence more valuable, by the confrontation of theoretical models with available statistical data. To do so is one of the main tasks of econometrics. In comparison to the magnitude of that task, work on it has hardly begun.

II

While thus being in wholehearted agreement with Orcutt's views as to the importance of isolating, and studying the effects of, instrumental variables, I find myself less confident, and partially in disagreement, with regard to the research proposals he bases on these views. I am now referring to his remarks about both the need for and the possibility of statistically testing the exogenous character of variables treated as such in econometric models. While I feel that our present knowledge and understanding of the issues in question is very incomplete, I venture to offer in tentative and unsystematic form some comments that would lead at least to problem formulations different from those chosen by Orcutt.

It will help first to have another look at the two main principles of classification of variables that have entered into the discussion: exogenous versus endogenous variables, and instrumental versus noninstrumental variables. The partial overlapping of these classifications can be visualized as follows:



All classifications in this alignment must be regarded as forward-looking. A variable is called instrumental if it is regarded as con-

trollable by the implementation of policy decisions, even though in some past period it may either have been left uncontrolled or may have been controlled in response to endogenous variables. If rain-making technology should develop sufficiently, it may make sense to construct models treating rainfall as instrumental, although in the past it has been an uncontrollable exogenous variable. It also makes sense to discuss business cycle policy with the help of a model treating the income tax rates as instrumental, even though these may in some past period have been set in response to experienced budget surpluses or deficits in an attempt to balance the budget. In this case the income tax rate, instrumental in a forward-looking sense, was endogenous in the past.

It may be remarked in passing that, while Orcutt does not seem to distinguish explicitly between controllable and uncontrollable exogenous variables, his remarks about the need for studying the continuity properties of economic time series apply properly to the uncontrollable exogenous variables. These variables, and the random disturbances, the policymaker must take as they come. Hence it will help him to know how much variability to expect of them in a given lapse of time.

The distinction between endogenous and exogenous variables has received further clarification in a recent study by Herbert A. Simon,[6] which will soon appear. He treats this dichotomy as a special case of the more general notion of a causal hierarchy of variables entering into a self-contained model, that is, into a model containing as many equations as non-instrumental variables. Without attempting here to anticipate his very illuminating and rigorous exposition, its essence may be summed up roughly in the statement that a variable z is classified as causally antecedent to a variable y if available policies that change z also change y, whereas other policies exist that change y but not z. Available policies are or can be represented by instrumental variables. (If these are to be included in the causal hierarchy, they come before all other variables and, if properly chosen, are coordinate rather than subordinate to each other.)

The relation of causal antecedence of y over z corresponds precisely to the distinction Orcutt draws between I and Y, respectively, in his example. The additional element in Simon's discussion is the explicit recognition that this hierarchy is a characteristic of the model utilized, of the policies the model recognizes as available, and of the manner in which the model represents the points of impact of these policies.

As an illustration consider the following example of a self-contained model of three equations with three instrumental variables p_1, p_2, p_3 and three non-instrumental variables y_1, y_2, z,

$$(a)\begin{cases}(a.1) & f_1(y_1, y_2, z, p_1) = 0\\ (a.2) & f_2(y_1, y_2, z, p_2) = 0\\ (a.3) & f_3(z, p_3) = 0.\end{cases}$$

The causal hierarchy gives z antecedence [7] over the pair y_1, y_2, because z is unaffected and y_1, y_2 are affected by changes in p_1, p_2; whereas all three variables y_1, y_2, z are affected by changes in p_3. The same example permits us to illustrate the distinction between exogenous and endogenous variables as a special case of causal hierarchy. We can say that equation (a.3) determines z, and that equations (a.1) and (a.2) together determine y_1 and y_2, given the causally antecedent variable z. For this reason, the submodel consisting of equations (a.1) and (a.2) is called a complete sub-model, and the variable z is called exogenous, the variables y_1, y_2 endogenous, *with respect to that sub-model.* This example shows that the assertion that certain variables entering into a certain complete (sub-) model are exogenous can be substantiated only by information about the form of equations outside that (sub-) model — "form" meaning here both the set of variables entering in and the policies impinging on these additional equations.

III

The foregoing statements, as well as most of Simon's analysis, apply to conceptual models of

[6] H. A. Simon, "Causal Ordering and Identifiability," ch. III of *Studies in Econometric Method*, ed. Wm. C. Hood and T. C. Koopmans, Cowles Commission Monograph No. 14, John Wiley & Sons, New York (to be published).

[7] To reach this conclusion, it must be assumed that the functions f_1, f_2, f_3 actually depend on the variables shown, that (a.3) can be solved uniquely for z whatever the value of p_3, and that (a.1) can be solved uniquely for y_1 and y_2 whatever the values of z, p_1, p_2.

the deterministic (non-stochastic) type, in which the path of each variable during a certain period of time is uniquely determined by the paths imposed on the instrumental variables (and by initial conditions, if some variables enter into the equations also with time-lags). Additional complications arise if the form of the model is no longer taken as a datum. Most of Orcutt's observations concern the problem of specification: the choice of the model, and the nature of the evidence that can be adduced in support of that choice. For purposes only of limiting the scope of the present discussion, let us still assume that all equations are regarded as indubitably linear. In this case the choice of the model consists in listing the variables; in subdividing the list into instrumental (controllable exogenous), uncontrollable exogenous, and endogenous variables; in listing the equations; and in specifying which variables enter into which equations, with what choice or variety of time lags.

The first remark I wish to make rests on what was said above about the forward-looking character of the concept of an instrumental variable. A variable which has not been controlled in the past is not for that reason necessarily uncontrollable. If two variables have not moved independently in the past, this does not necessarily preclude these variables from being used as two independent instrumental variables in the future. Congress may decide upon different tax rates for people over and under 40 years old, even though this has not been done so far. Hence, in attempting to answer the question whether or not a given list of instrumental variables is correct (which is logically a question of causal hierarchy), we shall when past experience is inconclusive have to use legal, institutional, and technological knowledge about the nature of the powers of government, the Federal Reserve System, or other agencies or groups of individuals whose policies are under discussion.

If this is accepted to be the type of evidence required for the choice of controllable exogenous variables, the second problem is that of distinguishing uncontrolled exogenous variables from the endogenous variables. There are several stages of complication in this problem. In order to take these piecemeal, let us first argue (unrealistically) as if the number of observable variables entering into a model can be held down to a finite and moderate number without introducing unobservable random disturbances. In such a linear non-stochastic model, given a moderate but sufficient number of observations, there is no problem of estimation beyond that of identifiability. Each equation can be determined exactly from the observations if the set of variables excluded from (not entering into) it is sufficiently numerous, and sufficiently different from the sets of variables excluded from other equations. An equation for which these conditions are not met[8] cannot be determined. Exclusions of variables beyond the minimum needed for identifiability (i.e., over-identifying specifications) are susceptible to testing (confirmation or refutation) from the observations. However, whether partly susceptible to testing or not, the information represented by the lists of variables excluded from each equation has no bearing whatever on the specification as to which variables are exogenous. This question of causal hierarchy can only be answered by information about the form of equations outside the model. Assurance that a given variable is exogenous can only be obtained by qualitative knowledge of the variables causally involved in *its* generation. If the model can be extended by additional equations describing the generation of the presumably exogenous variables, the needed information is of the same type as that required for identifiability: lists of variables occurring in the additional equations that make the model self-contained. Where the variables in question are often non-economic in character, the required knowledge may not at present be attainable by explicit extension of the model to cover a wider range of phenomena. The decision resorted to in such cases is called hypothesis by the econometrician and judgment by the economic statesman. The cost of misjudgment is obvious. Incorrect statements as to the effects of available policies are made if variables actu-

[8] For details of these conditions, and for a more inclusive (rank) condition depending also on the values of the coefficients of the other equations, see T. C. Koopmans, "Identification Problems in Economic Model Construction," *Econometrica*, April 1949, to be reprinted as ch. II, *Studies in Econometric Method*. The interrelations between causal ordering of variables and identifiability of equations are systematically explored by Simon, loc. cit.

ally endogenous are treated as exogenous. It is equally obvious that judgments of this type are made and need to be made continually by policy-makers.

Additional difficulties arise in the (more realistic) case in which observation does not extend to all relevant variables. Unobserved stochastic variables, that is, variables subject to a hypothetical probability distribution, are introduced to represent variables and equations not explicitly incorporated. In such stochastic models the concept of an exogenous variable must, of course, be defined afresh, and the definition that has been used[9] gives it a somewhat narrower meaning. By way of example, in the stochastic model analogous to (a) above (in which we do not explicitly show the instrumental variables)

$$(b)\begin{cases}(b.1) & g_1(y_1, y_2, z, u_1) = 0\\ (b.2) & g_2(y_1, y_2, z, u_2) = 0\\ (b.3) & g_3(z, u_3) = 0\end{cases}$$

the variable z has been called exogenous to the sub-model (b.1), (b.2) only if the stochastic variable u_3, and hence also z itself, is distributed independently of the stochastic variables u_1, u_2. The specification of exogeneity according to this "stochastic" definition must frankly be recognized as an approximative device whereby a "complete" sub-model is obtained, for which statistical methods of estimation and of hypothesis-testing can be and have been developed. I believe that Klein's choices of exogenous variables cited by Orcutt should be interpreted in this way.

Orcutt asks for a statistical test of the exogeneity of z in this stochastic sense, I presume. Before going into the question of the chances for such a test to be sucessful, let me say that, in comparison with the non-stochastic case, the cost of incorrectly specifying exogeneity has increased. Besides the incorrect assessment of policy effects from numerically correct coefficients, noted already in discussing non-stochastic models, we now have the inconsistency of estimation of the coefficients, inherent in an erroneous assumption of exogeneity, which is bound further to distort the assessment of policy effects.

IV

The foregoing discussion leaves no doubt that it would be very important to have a test of exogeneity in the stochastic sense, a test that has some power of discrimination. Unfortunately, if I may venture a conjecture, it does not seem to me that the chances are good for such a test to be really informative, if applied to actual data. The difficulty lies in the necessity, in all statistical testing, to specify a set of maintained (unquestioned) hypotheses. These maintained hypotheses are not themselves subjected to test but indicate the range of alternative hypotheses held possible if the one tested is untrue. For a test of the hypothesis that a parameter has a certain value to be possible, the maintained hypotheses must be sufficiently strong to make that parameter identifiable. We can only construct tests of overidentifying (more generally: observationally restrictive[10]) hypotheses.

Now, in relation to the maintained hypotheses of some of the simpler models, a priori specification as to which variables are exogenous (in the stochastic sense) is required to obtain identifiability of the coefficients of these variables. This specification then escapes all possibility of a test. With some ingenuity, it is often possible to "set the stage" for testing a certain hypothesis, by introducing sufficiently strong maintained hypotheses into the model so that, in conjunction with these, the hypothesis in question has observable implications. In the model (b), a specification that the distribution of u_1 and u_2 is the same at successive points or periods in time (at which observations are made) could be used for that purpose.[11] While we often use that specification (even narrowed down further to a normal distribution of u_1, u_2) to suggest estimation formulas in cases where

[9] T. C. Koopmans, "When Is an Equation System Complete for Statistical Purposes?," ch. XVII, *Statistical Inference in Dynamic Economic Models*, Cowles Commission Monograph No. 10 (New York, 1950).

[10] For definitions of these concepts, see sec. 2.6 of T. C. Koopmans and Olav Reiersøl, "The Identification of Structural Characteristics," *Annals of Mathematical Statistics*, 1950, pp. 165–81, reprinted in Cowles Commission Papers, New Series, No. 39.

[11] See A. Wald, "Remarks on the Estimation of Unknown Parameters in Incomplete Systems of Equations," ch. VIII, *Statistical Inference in Dynamic Economic Models*.

the consistency of the estimates does not depend on it,[12] it is harder to justify using the same specification in cases where without it consistent estimation is impossible. While therefore one cannot categorically declare a test of exogeneity (by the "stochastic" definition) to be impossible, it does seem to me that, with models such as those used so far for actual estimation purposes, the prospects for obtaining a clear verdict from the available number of observations are not encouraging. I would greatly welcome being proved wrong in this estimate of the situation.[13]

If no promising tests of exogeneity are found, the task remains of assessing the limits of error inherent in policy conclusions drawn in a state of uncertainty with regard to the exogenous character of certain variables. This is a special case of the broader problem of specification error and the strategy of model construction.[14] If doubt remains about a basic specification not subject to conclusive test, the only remaining line of defense is a study of the effect on policy conclusions of presumably possible degrees of departure from the specification in question.

To sum up, it is believed that instrumental variables must and can be recognized from general legal, institutional, and technological knowledge as distinct from statistical observations on economic time series; that (as urged by Orcutt) further study of the variability and continuity properties of uncontrollable exogenous variables and random disturbances is needed; that the evidence on which the choice of exogenous variables rests must be sought primarily in qualitative knowledge about the place of the variables in question in the causal hierarchy, with slight chances of corroboration from statistical tests utilizing time series; and that where this evidence is insufficient studies of the effects of erroneously specifying exogeneity are needed.

[12] See H. Chernoff and H. Rubin, "Asymptotic Properties of Limited-Information Estimates under Generalized Conditions," ch. VII, *Studies in Econometric Method* (forthcoming).

[13] The above remarks are addressed to the possibility of the maintained hypotheses being made strong enough to permit a test of exogeneity of some power. If the answer were to be in the affirmative, difficult problems in the technique of test construction would be encountered next. The exogeneity assumption has the role of "closing" the model to make it into a complete model. The maintained hypothesis must therefore leave open the possibility that the model is incomplete, a situation not envisaged in the current theory of testing hypotheses.

[14] For a discussion of problems of this type, see L. Hurwicz, "Some Specification Problems and Applications to Econometric Models," Abstract, *Econometrica*, 1951, pp. 343–44.

II

By J. Tinbergen

I am in almost complete agreement with Dr. Orcutt's views; in fact it is striking to what an extent my own forthcoming publication "On the Theory of Economic Policy" is based on the same approach. It should be recognized, however, I think, that Frisch, in his "Memorandum on Price-Wage-Tax-Subsidy Policies" [1] laid the foundations to this approach.

First of all, I want to join Dr. Orcutt's contention that prediction is not such an essential activity of the economist or the econometrician; the future course of any economic variable always contains random components which we do not know and which make our prediction much less accurate than the solutions of "variation problems." By the latter I understand the problem of indicating how a certain variable changes if one of the data is varied, in particular one of the "political parameters."

The specification of the variables chosen as exogenous, on which Dr. Orcutt lays so much emphasis is, of course, very important. In principle, it should be based, in my opinion, on a priori rather than on statistical considerations. Generally speaking the exogenous variables are either non-economic or outside the market system studied. It is true that it is only by hypothesis that these variables do influence the endogenous variables without being themselves influenced by them; and it should be admitted that the testing of this hypothesis is useful and necessary. It is also true that in recent econometric and economic analysis certain variables are somewhat too easily assumed to be exogenous. Keynesian economics is not without guilt here; in particular the as-

[1] R. Frisch, "A Memorandum on Price-Wage-Tax-Subsidy Problems as Instruments in Maintaining Optimal Employment," The University Institute of Economics, Oslo, published as a U.N. Document, April 1949.

sumption, often made, that investment and public expenditure would be completely exogenous goes too far. As an alternative approach my own treatment of investment as an endogenous variable may be mentioned. The solution should be based on the distinction, for such variables, between an "autonomous" and a "dependent" component; a distinction used by many authors (Frisch, Meade) under varying names.

The problem whether certain exogenous variables are interrelated is only another aspect of the same subject. Certainly some authors have been fully aware of a possible interdependency. In a study on the influence of wages on employment, where wages were supposed to be exogenous, De Wolff and I used as one of the other exogenous variables productivity. For the short-term impact of wages on employment this seems legitimate; when investigating the long-term influence we added the hypothesis that productivity would vary in dependence on wages.[2]

Statistical testing of hypotheses on the correlation or lack of correlation between certain exogenous variables should be, in principle, based on a theory of the movements of these variables, which may either be a complement, for the outside economy, to the economic theory of our model, or a non-economic theory of the behavior of investors or of government, or both. Oversimplified devices such as looking at the observable correlation between the exogenous variables seem somewhat dangerous. Often a priori information may be more reliable. A complete theory, as just indicated, deserves our preference, however.

III

By Nicholas Georgescu-Roegen

The title of Mr. Orcutt's paper can hardly pass unnoticed. But whether it will arouse enthusiasm — this could be almost unlimited were it not for the qualifying term "partial" — or whether it will be met with deep skepticism is rather difficult to say. The writer's best guess is that the attentive reader, even if his anticipations were full of optimism, will ultimately wind up with a feeling of dissatisfaction. Indeed, Mr. Orcutt's Pegasus, more impatient than the mythological one, did not wait for the death of the econometric Medusa. As if realizing that the birth of his Pegasus prior to this death may be the cause of the failure to strike a new source of inspiration, Mr. Orcutt, hoping to put things in order, tries — if not to finish off the Medusa — at least to speed up the ceremony of her sacrifice.

In the introductory paragraph and first two sections of his paper, Mr. Orcutt maintains the spirit of the optimistic reader at a high pitch. Here we find formulated one by one all the dreams of an econometrician and, more especially, of a policy-maker. (Mr. Orcutt apparently thinks that the services of the economist *per se* are not worthy of sharing, even in a very modest way, the glory of collaborating to solve the problem.) While a careful listing is made of all the things which, in Mr. Orcutt's opinion, would make the life of a policy-maker very comfortable — but which also would deprive him of an inestimable glory whenever successful — the grounds for the case against econometrics are gradually built up.[1] As a start in this direction, we are told that the "econometricians have failed to attack in any force problems whose solutions not only would be useful to policy-makers but whose solutions may be more feasible," and that they have spent their time playing around with other problems whose solutions are, in Mr. Orcutt's opinion, either not feasible or less feasible. With such a horrible report-card, the econometrician is treated with the usual fatherly advice: first, that "more emphasis needs to be placed on building and testing models" and, second, that "more study of the continuity properties of economic time series is thus needed." It is not difficult to guess that the laggard schoolboy would have preferred both to have been spared the admonition and to be actually helped with some of his homework. In the latter connection, Mr. Orcutt is apparently not open to criticism since in the next two sections (III and

[1] In his enthusiasm, Mr. Orcutt is ready to widen his target to include all physical sciences, challenging their success in the operation of control systems. This point will not be taken up in the present paper.

[2] J. Tinbergen and P. de Wolff, "A Simplified Model of the Causation of Technological Unemployment," *Econometrica*, 7 (1939), 193, in particular p. 205 (middle).

IV) he develops a system which, in his opinion, could help the econometricians improve their grades.

One should normally go directly to the central part of Mr. Orcutt's argument and, in a written criticism, omit discussion of the earlier sections on the ground that they are introductory in character. However, some of the difficulties connected with the theme of the paper are already exhibited in the preliminary remarks. This is why one may be justified in exploring them before proceeding further.

In contrast with his strictness regarding the definitions used by other econometricians, Mr. Orcutt frequently leaves the reader in confusion as a result of the imprecision of some of his terms. More space could have been profitably diverted from obvious generalities to an explanation of what the author understands by various terms, for instance, by "more feasible solutions" and, more especially, by the "instrument by which the policy-maker may modify the course of the actual." To the very end of the paper, the reader cannot find out whether Mr. Orcutt by "instrument," (becoming successively "instrument of adjustment" and "control instrument"), means:

(a) the institutional means at the disposal of the policy-maker (e.g., the power to change the tariff, or to introduce rationing, etc.);
(b) the variables or the parameters which can be modified by the measures mentioned in (a);
(c) the theoretical (or econometric) relation existing between some variables (b) and other variables of the economic system;
(d) the actual effect of a change of the variables (b) on the other variables of the system;
(e) any other concept which Mr. Orcutt might have had in mind.

This ambiguity certainly does not help us reach a clear-cut picture of the equipment necessary to the policy-maker and the method of using it, as seen by Mr. Orcutt. And this is the origin of a haze which extends over the main part of the argument, which also considers the tools useful for the policy-maker.

The last paragraph of section I offers the first sign of Mr. Orcutt's opinion that the economic factors which "we know how to control and that we contemplate using for control purposes" are normally found among "variables" and not among "parameters." "There has been some tendency," we read, "to think of many policy actions as consisting of changes in the parameters of the econometric models," but Mr. Orcutt regards this tendency as groundless since the parameters cannot be altered without having "one or more auxiliary models [relating] the parameter values" to the controllable variables.[2] We do not know exactly what Mr. Orcutt means by "parameters of a model," but it is clear that under the most widely accepted use of the term — that which stems originally from multiple or general equilibrium models [3] — we find some parameters among the oldest and the most preferred channels for carrying out economic policies. The outstanding examples of this are the "tariff schedule" and the "tax schedule." Furthermore, the distinction between controllable and uncontrollable factors cannot be made in the abstract, independently of the problem at hand, or by an a priori formal approach, as is the usual definition of the parameters.[4] It is not possible to know — without the help of economic analysis and its great ally, economic history — which factors are controllable and which are not. They may be parameters as well as variables, and the econometrician cannot alter their quality. He has to accept them as such and build his models accordingly.

One may heartily join Mr. Orcutt in wishing

[2] The term "control" is loosely used by Mr. Orcutt. At times, it refers to variables "we *wish* to control" — i.e., the ultimate objective of the policy-maker — while, at others, it is connected with those variables "that we contemplate *using for control purposes*" — i.e., the factors over which the policy-maker has direct control. (In both quotations, italics have been added.) It is in the latter sense that the terms "control" and "controllable" are used above by the present writer.

[3] Cf. Jacob Marschak, "Statistical Inference in Economics: An Introduction," in *Statistical Inference in Dynamic Models* (Cowles Commission Monograph, No. 10, ed. T. C. Koopmans), pp. 7–8.

[4] See, however, T. Haavelmo, "The Probability Approach in Econometrics," *Econometrica*, 12 (Supplement, July 1944), 3.

that more "models which include as exogenous variables those variables that we know how to control" be built and tested, but wishing alone will not help.

The main problem of building models is rather that of making them complete from the point of view of both economic theory and statistics.[5] If this criterion is followed, the group of exogenous variables cannot be arbitrarily set, neither can the subgroup of controllable exogenous variables be so set. In each particular case they are determined by the structure of the problem under study.[6] And this is why the same variable may be endogenous in one case, and exogenous in another. The decisions regarding the specification of exogenous variables are made, therefore, so that the problem may be handled in the best way and not, as Mr. Orcutt states, so as to "arbitrarily set[ting] the limits of the problem under consideration."[7] In order to make sense of each particular model this is the logical procedure. Thus under certain circumstances, a Leontief model open with respect to "households" may be used, while, under others, the same model may be open to "other countries." The first would be justified under the assumption of some type of rationing, which in turn would justify the consideration of the "bill of goods" as exogenous; the second, under the assumption of foreign trade control, which may make the "exports" exogenous. Changes of this kind may prove to be at times very useful for analytical purposes.[8] They have nothing to do with models aimed at helping the policy-maker.

In section III, as the argument draws closer to Mr. Orcutt's theme — which centers upon the concept of the exogenous variable — a new accusation, far more serious than any previous one, is thrown at "economic theory (*sic*) and econometrics." This time it is that their "literature is far from explicit about the difference between endogenous and exogenous variables." This accusation too is not supported by any evidence.

However defective the literature may be on this point, from it Mr. Orcutt gathered a definition of exogenous variables which he seems to adopt temporarily. According to this definition, the "exogenous variables are (those) which affect the economic system but are not in turn affected by it, *or at least are only affected to a negligible degree by it.*"[9] This definition strikes a new chord, capable of deep and multiple resonances, and would have induced the writer to consider it at a great length, had this not already been done by some of the best contributors to the theory of statistical inference.[10]

Abstracting for the time being from the italicized part and taking the remainder of the above definition *ad literam*, no exogenous variables can be of an economic nature. Neither can they be "sociological, political, and psychological factors,"[11] nor factors describing the state of the arts and geographical location. If we include as endogenous all these variables, we are left with a system "open" only with respect to the initial cosmological conditions: time, the inalterable properties of matter, and its initial distribution in the space. Such a classification shades into an almost metaphysical scheme and loses all importance for any practical inquiry. It is the inclusion of the italicized phrase — a weakening condition — which makes the concept of exogenous variables useful for econometric analysis and, indeed,

[5] See the penetrating analysis of T. C. Koopmans, "When Is a System Complete for Statistical Purposes?" (Cowles Commission Monograph, No. 10), pp. 393–409.

[6] The considerations that enter into the logical process which is involved here are basically those provided, as mentioned before, by economic analysis and economic history. They are, in Mr. Orcutt's opinion, only "some a priori knowledge of unspecified source."

[7] The present writer does not deny that a few isolated cases may be perhaps found where a certain degree of *irrationality* is present in the particular treatment of a problem, but only asserts that, if these cases exist, they do not constitute the general rule. Besides, an entire discipline — such as is the object of Mr. Orcutt's attack — cannot be made responsible for the errors committed by one of its users. One single specific example would have helped the reader to make better use of Mr. Orcutt's criticism. At the beginning of section III, he mentions *by title* some of the best known works, those of J. Tinbergen, L. R. Klein, and Colin Clark, and at no place in his paper is a connection established between his criticism and the methods used by these authors. To be more explicit, the reference by title to their works may very well be left out: the argument of Mr. Orcutt will be *in no way* affected by the omission.

[8] Koopmans, op. cit., p. 394.

[9] Italics added.

[10] E.g., Marschak, op. cit., p. 8; Koopmans, op. cit., pp. 393 ff. The above definition is in fact what Koopmans calls "the causal principle" (*ibid.*, p. 394).

[11] Koopmans, op. cit., p. 402.

for all social sciences. It seems, therefore, a natural thing to formulate the definition of the exogenous variable on the basis of the weak, rather than on the strong, causal principle. For the sake of greater rigor, one should also try to make more precise the meaning of "negligible degree." Here we are confronted again with the difficulty of giving a definition to a loose concept so that it make sense as an analytical tool. Some influences work their effect quickly, others more slowly. In some cases, the effects are of a more lasting nature, almost cumulative in character, in others, the effects are short-lived, leaving no trace. Furthermore, the limit of "negligible degree" cannot be set uniformly.[12] The setting of a reasonable limit must be left to the model-builder or to its user. The exogenous variable is a relative and loose concept,[13] and very little indeed can be done about it.[14] Almost any model will provide a good illustration of this point. One may start with the three Marshallian models, the market, the short-run, and the long-run equilibria, characterized by three different exogenous variables, which are respectively the day supply, the size of the capital equipment, the state of the arts, and by one common to all, the tastes of a stationary population. Again, price constellations exercise their influence upon the quantities produced as well as upon the state of the arts. It is only because it takes longer for the latter than for the former to show a visible effect of this influence that treating the input-output coefficients as exogenous variables in a Leontief model is justified.[15]

The present writer is at a loss, therefore, to find a reasonable justification for Mr. Orcutt's repetitious complaints, some of which leave no room for a less strong interpretation. This is the case in the italicized statement that, "in any case, *the specification* [of which variables are exogenous] *is not subject to any test whatsoever.*" Later on, the econometricians face their n-th accusation, that they have "almost complete[ly] neglect[ed]" the "testing [of] hypotheses about which variables are wholly or partially (*sic*) exogenous to the economic system." But this time, Mr. Orcutt had consumed exactly one half of his paper in fighting econometrics, apparently for the sole purpose of preparing the ground for the "redirection." He then decides to "have another look at the definition of an exogenous variable and see what its definition means in statistical terms."

Despite the fact that the definition thus far used by Mr. Orcutt for the exogenous variable is by no means stochastic in character, no preparation is made for the turn toward statistics. The latter comes therefore as a surprise, which, however, is not to be the last.

In order to illustrate his main theme Mr. Orcutt uses a very simple, linear model (1), whose definition is given piecemeal. When the reader finally makes it out from different bits of information scattered throughout the argument, he finds, not without surprise, that Mr. Orcutt's new definition of exogenous variables is a *sui generis* interpretation of the stochastic model used by Professor Koopmans in his paper, already quoted above. Indeed, Mr. Orcutt's model consists of two sets of equations

$$(A) \quad (A_1) \quad Y - a - bI = 0 \quad (A_2) \quad I = 0$$

which determine the solution of the entire system.[16]

One point needs here a special emphasis. The definition of the exogenous variable is only interpreted by the structure of the system (A) and is not equivalent to (A). One should add to the system (A) the condition that the first equation cannot be used to explain the values

[12] The complete failure of a similar attempt, that of the classical statisticians, to set uniformly the limit of the significance of probability at $P = .05$, is very instructive in this regard.

[13] If it were otherwise, we would be in a position to answer the question that comes frequently from our sophomore classes: how long is the short run and how short is the long run?

[14] Of course, we can formalize its definition, as Professors Marschak and Koopmans did, but, while this helps tremendously to clarify our ideas and to treat some important problems of statistical inference, its use in econometrics is to help us in building the model so as not to contradict ***known*** facts about the nature of variables — supported by economic history, economic theory and, at times, by other tests — rather than in discovering *new* facts about these variables. *Infra*, fn. 25.

[15] For other examples, see Koopmans, op. cit., p. 394.

[16] Cf. Marschak, op. cit., p. 8. The above equations correspond respectively to the systems (1.4) and (1.2). This writer feels that by confining the argument to two variables instead of two systems many fine points of the problem are obscured. It is because of this limitation that we arrive at expressions such as "I is exogenous to Y."

of I, were the second equation to be hidden from us.[17] It is this qualitative condition that completes the definition of the exogenous variable. Whatever follows from now on is built on top of the concept. This is not, however, Mr. Orcutt's opinion, since he tries to define the exogenous variable with the help of properties other than those just mentioned.

The problem that comes naturally next is to see how we can make use of the model (A) in order to interpret a body of observed values of Y and I. For this purpose, a stochastic scheme must be introduced in (A). This can be done in various ways which, for the purpose of exposition, can be exemplified by introducing:

(a) errors in the observed values of the variables;
(b) shocks suffered by the theoretical relations (such as parallel shifts in a straight-line demand);
(c) other stochastic influences (such as changes in the slope of a straight-line demand).[18]

Mr. Orcutt chooses to consider the "simple shock model," that is, (b). This means that during the period of observation, the straight lines (A1) and (A2) suffer shocks, ϵ and η (Chart 1).[19] The true position is E_0, and the observed is e. The system (A) becomes

$$(B) \quad (B1) \quad Y - a - bI = \epsilon \quad (B2) \quad I = \eta$$

These are particular cases of equations (3a) and (3b) used by Professor Koopmans.[20] Further, following the same author, we introduce the assumption that the random variables ϵ and η are independently distributed. Thus, if $f(\epsilon, \eta)d\epsilon\, d\eta$ is the distribution of ϵ and η, then [21]

$$(C) \qquad f(\epsilon, \eta) = f_1(\epsilon)\, f_2(\eta).$$

This is introduced with the idea of keeping the two *systems* (A1) and (A2), or (B1) and (B2), free from any possible connection other than that attributed to the exogenous variables.

Because of (B2), the relation (C) becomes

$$(C') \qquad f(\epsilon, I) = f_1(\epsilon)\, f_2(I)$$

which brings us to Mr. Orcutt's definition that the exogenous variables are those which "are distributed independently of the excluded variables or shock terms as they are sometimes called." [22] If the term "shocks" is to be used here unambiguously, one should make clear that here only the shocks of the relations involving the endogenous variables are considered.[23]

Mr. Orcutt differs from Professor Koopmans in the fact that the former thinks that "to say that I and ϵ are independently distributed would be equivalent to saying that of the included variables I is the exogenous one." Apparently, Mr. Orcutt does not feel that this equivalence requires a proof since he does not offer one. It is not difficult, however, to see the weakness of this point which, unfortunately, serves as pivot for the positive theme of Mr. Orcutt. Indeed, it is only in the light of system (A) and of the additional specification of the direction in which causality works that the condition (C') acquires a meaning. The latter

[17] E.g., if I were the rainfall and Y the crop-yield, the first equation could not offer an explanation of the rainfall in terms of the crop-yield. (It could offer, however, a method of an a posteriori estimation of the rainfall, but this is another problem.) In a note ("A Suggestion for Notation in Mathematical Economics," *Quarterly Journal of Economics*, LIV, November 1939, 165–67), Andrew W. Edson made the interesting suggestion that, in economic mathematical relations we should use an arrow, in addition to the equality sign, in order to specify the "direction" of the causal interpretation of such a relation. In other words, the arrow will show which side of the relation contains the endogenous variable.

[18] Cf. Marschak, op. cit., pp. 20–21.

[19] The relation (A2) might have been written $I - I_0 = 0$, without altering the argument.

[20] Op. cit., p. 395.

[21] *Ibid.*, p. 396, relation (4).

[22] This statement justifies completely the interpretation of Mr. Orcutt's ϵ as a *shock* term and the presentation of his model as was done above. Another alternative, available here only because of the simple structure of the model (1), would have been to interpret ϵ as determined by an *error* of Y and of I. The reader had to wait until the end of the argument to learn which interpretation was used by Mr. Orcutt.

[23] It is easy to see that from Professor Koopmans' basic assumptions regarding his shock model, i.e., from assumptions 3.1, 3.2, and 3.3, it follows immediately that the distribution of the exogenous variables $(x_{G+1}, x_{G+2}, \cdots x_N)$ is independent of that of the shocks $(u_1, u_2, \cdots u_G)$ of the relations involving the endogenous variables $(x_1, x_2, \cdots x_G)$. If $\chi(u_1, u_2, \cdots u_G, u_{G+1}, \cdots x_N)$ is the distribution density, then $\chi = f_1(u_1, u_2, \cdots u_G) \times f_2(u_{G+1}, \cdots u_n) \dfrac{\delta(u_{G+1}, \cdots u_n)}{\delta(x_{G+1}, \cdots x_n)}$

supplements the former, but is *not* equivalent to it.[24]

Neither can Mr. Orcutt claim that, since the condition (C′) is not a definition of exogenous variables, it constitutes at least a partial test which all such variables must fulfill. Indeed, (C′) has no meaning outside the simple shock models. Thus, the only thing that remains from the theme developed in section IV is that the distribution of the exogenous variables must — in the case of simple shock models — be stochastically independent of that of the

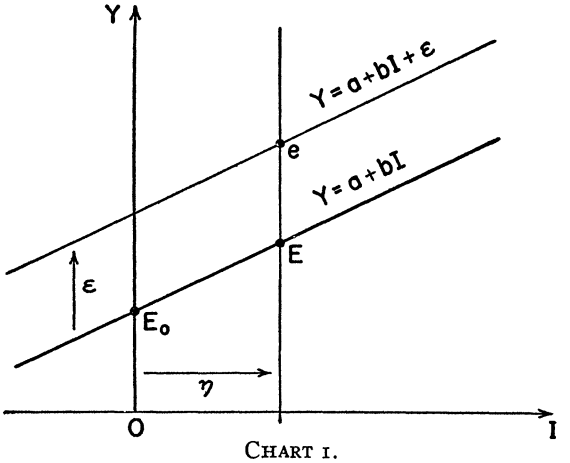


CHART 1.

shocks. As a consequence of this principle, we are told that "if this correlation [between I and ϵ] turned out to be significantly different from zero, then some modification of the model would seem to be called for." But Mr. Orcutt is too experienced a statistician to be unaware of the ensuing difficulty. The problem, as it stands before the econometrician, is not only to determine which variable — national income or investment — is exogenous, but to determine also a and b. And for this task, the only available information about the reality under study is a scatter formed by a number of points, such as e (Chart 1). And as Mr. Orcutt rightly observes, "unfortunately, the usual fitting process ensures a selection of values for a and b such that the correlation between the variable selected to be the independent one and the obtained residuals must necessarily be exactly zero." But this is not all, however. Had the econometrician selected the national income as the independent variable, *the same scatter would have confirmed his choice of Y as the endogenous variable*![25] While this situation may help the econometrician to maintain cordial relations with both the Keynesians and the anti-Keynesians, it reveals that, without a model accepted prior to the study of the statistical inference, the econometrician cannot solve any problem. On the other hand, the statistical inference will confirm the choice of any model provided the latter follows, be it only in general lines, the map of data. To use an analogy, an old favorite of the present writer, the situation of the statistician is such that he stands very little chance of disproving that in every log there is a beautifully sculptured Madonna — simply because his tools are such that when he tries to get inside a piece of wood, by this very procedure, he carves the statue. Mr. Orcutt promises to put an end to this tragic fate of the modern Midas, with a subsequent paper devoted to the choice of independent variables. Until then, the econometrician is again invited, in a formal finale, to double and redouble his efforts at studying.

IV

Summary Comments by Guy H. Orcutt

I am of course pleased to find myself in such close agreement with Tinbergen, whom we all recognize as one of the truly great econometricians. I am also grateful for his references to the work of Frisch and to his own study with P. de Wolff.

The question I might raise in connection with Tinbergen's remarks is in his reliance on a priori knowledge. To me a priori knowledge is just knowledge that has been gleaned from previous studies. It is of course reasonable to hold that each study should build on what has gone before and in this sense to make use of a priori knowledge. But to label some piece of knowledge as being given a priori can hardly add any support to its validity over and above

[24] Using Mr. Edson's suggestion (*supra*, fn. 17), the line connecting I and Y in Mr. Orcutt's triangular diagram should be replaced by an arrow pointing toward Y. It is the logic which justifies the direction of the arrow — the causal principle — and not the type of stochastic relation between ϵ and I that determines which variable is the exogenous one. As a matter of fact, ϵ is introduced in the diagram for stochastic, and not for fundamental, reasons.

[25] *Supra*, fn. 14.

the evidence which has been used to establish it. We still must somehow learn from experience.

I find myself in rather close agreement with Koopmans' remarks except on a few points. Thus, we agree on the importance of explicit incorporation into econometric models of instrumental variables and of intensification of the study of the continuity of economic time series. I would not agree with Koopmans that we can study with advantage the continuity properties of *only* the uncontrollable exogenous variables. This would seem to presume the untenable position that any use of any instrumental variable would drastically modify all the useful continuity properties of all of the endogenous economic variables. I regard as useful Koopmans' remarks about the two main principles of classification of variables that have entered into our discussion, but I could not fully accept the way in which he relates these classifications. In particular, to classify a variable as instrumental is not at all inconsistent with classifying it as endogenous. It all depends upon whether the instrumental variable moves or is made to move in response to the variation of other endogenous variables in the system. I agree that, "It also makes sense to discuss business cycle policy with the help of a model treating the income tax rates as instrumental, even though these in some past period have been set in response to experienced budget surpluses or deficits in an attempt to balance the budget." I would, however, want to go on and add that to find out the effect of exogenous variations of the tax rate one would indeed have to find some historical evidence of what happened when tax rates or some analogous variable did in fact vary exogenously. Much of what Koopmans has to say about the choice of controllable exogenous variables would seem to be appropriate only to the choice of controllable (i.e., instrumental) variables. The problem of discovering which variables can be controlled is indeed important. But this in itself does not by any means justify us in estimating, from historical data, their impact on other variables as though they had behaved as exogenous variables.

Koopmans and I are also in close agreement upon the great importance of correctly specifying which variables should be treated as exogenous. I think, although perhaps wrongly, that we are also in agreement that not only must some type of evidence based upon experience of some sort be depended upon to determine the specification, but that to date little if any explicit evidence on this point of any sort has found its way into published econometric models. At least this was one of my "repetitious complaints," to borrow Georgescu-Roegen's comment, and Koopmans certainly does not deny its validity. Koopmans, Simon, and myself all appear to be in fairly close agreement about the distinction between endogenous and exogenous variables and about the more general notion of a causal hierarchy of variables. However, Koopmans and I apparently are not in complete agreement about the ways in which causal structure may be tested. This is too big a subject to take up in these comments, but I hope that my position will be made clear in a forthcoming paper, "Actions, Consequences, and Causal Relations," which is to appear in the next issue of this REVIEW.

I could not make up my mind whether Georgescu-Roegen is accusing me of stealing Koopmans' ideas on exogenous variables and causal structures or of merely putting forth, in my abysmal ignorance, a poor likeness of them. Having written reviews of two major Cowles Commission books, including the Cowles Commission Monograph No. 10 referred to by Georgescu-Roegen, and having written three articles which bear directly on Cowles Commission procedures, I am in a poor position to claim ignorance. I have a very high regard for the works of Koopmans and Simon on this subject, but for the record I would like to point out that I certainly am not guilty of borrowing without giving credit since my own ideas on this subject were made public in my paper, "The Inference of Causation," which was read before the Econometric Meeting at Harvard, September 1950. An abstract of this paper is in *Econometrica*, 19 (January 1951), page 60.

The argument of part IV of my paper has not been challenged by either Koopmans or Georgescu-Roegen and, to the best of my knowledge, it is very pertinent to the efforts of econometricians. Most of Georgescu-Roegen's substantive comments have to do with the

definition, used in my paper, of exogenous variables and with a development of some of his own ideas about the meaning and specification of exogenous variables or more generally of causal structure. His ideas are interesting, although, if I understand him, I would reject his general position which would seem to make the specification of causal structure primarily a matter of theoretical convenience in completing the specification of models. In any case, this is not the place for a discussion of this topic. My point was that the specification of exogenous variables, in the sense meant by Koopmans and myself, is critical not only from the standpoint of estimation but also from the standpoint of policy implications. This being the case, it seems obvious that the specification should be supported by evidence. I further went on to assert that in studies of the type referred to this had not been done and that econometricians should exert themselves to produce such evidence. Georgescu-Roegen does not contradict me on any of this but merely says that it is the function of theory and history to provide such specification. This is fine but hardly removes the need for presentation of evidence. I would have thought it fairly well agreed that we must base all inferences on history (i.e., past experience). I would also regard it as obvious that in drawing inferences we take account of what we already think we have learned from experience (i.e., theory). And certainly it is true that history is the raw material of the statistician.

I am sorry that Georgescu-Roegen failed to see any positive content in any of my suggested lines of research. In any case, in trying to prove his point he might have tried to avoid rather consistently quoting me out of context. Thus, for example, my statement that "more emphasis needs to be placed on building and testing models or components of models which include as exogenous variables those variables that we know how to control and that we contemplate using for control purposes" becomes, when quoted and emptied of its entire meaning by Georgescu-Roegen, just "more emphasis needs to be placed on building and testing models." A little further insight into the way in which Georgescu-Roegen uses quotations may be gained by comparing the first sentence of his sixth paragraph with the last paragraph of my section I and from which he claims to be quoting. I can find the words and phrases quoted but the meaning seems to have undergone a slight transformation.

[Footnotes]

[4] **The Use of Econometric Models as a Guide to Economic Policy**
Lawrence R. Klein
Econometrica, Vol. 15, No. 2. (Apr., 1947), pp. 111-151.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28194704%2915%3A2%3C111%3ATUOEMA%3E2.0.CO%3B2-G

[5] **On the Theory of Business-Cycle Control**
J. Tinbergen
Econometrica, Vol. 6, No. 1. (Jan., 1938), pp. 22-39.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28193801%296%3A1%3C22%3AOTTOBC%3E2.0.CO%3B2-K

[10] **The Identification of Structural Characteristics**
T. C. Koopmans; O. Reiersol
The Annals of Mathematical Statistics, Vol. 21, No. 2. (Jun., 1950), pp. 165-181.
Stable URL:
http://links.jstor.org/sici?sici=0003-4851%28195006%2921%3A2%3C165%3ATIOSC%3E2.0.CO%3B2-A

[2] **A Simplified Model of the Causation of Technological Unemployment**
J. Tinbergen; P. de Wolff
Econometrica, Vol. 7, No. 3. (Jul., 1939), pp. 193-207.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28193907%297%3A3%3C193%3AASMOTC%3E2.0.CO%3B2-O

LINKED CITATIONS

[4] **The Probability Approach in Econometrics**
Trygve Haavelmo
Econometrica, Vol. 12, Supplement. (Jul., 1944), pp. iii-vi+1-115.
Stable URL:
http://links.jstor.org/sici?sici=0012-9682%28194407%2912%3Ciii%3ATPAIE%3E2.0.CO%3B2-F