MEMORANDUM

FRA

UNIVERSITETETS SOCIALØKONOMISKE INSTITUTT

OSLO

MJ. 6 November 1948.

AUTONOMY OF ECONOMIC RELATIONS

By

Ragnar Frisch

Trygve Haavelmo

T.C. Koopmans

J. Tinbergen

MEMORANDUM

FRA

UNIVERSITETETS SOCIALØKONOMISKE INSTITUTT

OSLO

MJ. 6 November 1948.

A U T O N O M Y O F E C O N O M I C R E L A T I O N S

By

R a g n a r F r i s c h

T r y g v e H a a v e l m o

T.C. K o o p m a n s

J. T i n b e r g e n

60.000.6.52.O.F.

Preface

For any economic relations introduced into our analysis we may ask: How autonomous is it? This question is extremely important. In one sense it is the most basic question one may raise in all sorts of econometric work. The Institute has found it expedient to reproduce in the form of a stencil-memo some contributions that have been made towards a discussion of this question, namely:

1. League of Nations. "Statistical versus Theoretical Relations in Economic Macrodynamics."

A memorandum prepared by Prof. Frisch for the Business Cycle Conference at Cambridge, England, July 18th - 20th, 1938, to discuss Professor J. Tinbergen's publications of 1938, for the League of Nations.

2. League of Nations. "Mr. Tinbergen's reply to Professor Frisch's note on "Statistical versus Theoretical Relations in Economic Macrodynamics."

3. "Probability Approach in Econometrics" by Trygve Haavelmo. Cowles Commission Papers. New Series, No. 4. 1944. Section 8 of Chapter II.

4. Parts of Professor Frisch's article "Repercussion Studies at Oslo." American Economic Review. Volume XXXVIII, No.3, June, 1948.

5. "Identification Problems in Economic Model Construction" by T.C. Koopmans. Cowles Commission Discussion Papers, Statistics: 316. July 19, 1948.

We are particularly glad to be able to reproduce the important contributions from Professors Tinbergen and Koopmans, both outstanding in the field of econometrics and both friends of the Oslo Institute. At one time, Professor Koopmans spent considerable time as a visiting research associate at the Oslo Institute. He is now Director of Research of the Cowles Commission, University of Chicago.

As Professor Trygve Haavelmo is "one of our own" it would not be fitting for me on this occasion to expound on the importance of the contributions he has made in this field.

In the field of statistical methodology the problem before us has given raise to what is now called Confluence analysis (Frisch, Koopmans, Reiersøl, Geary and others) and the related methods of Jerzy Neuman (Econometrica Jan. 1948). These more technical aspects of the problem are not covered in the present excerpts.

Ragnar Frisch.

1. League of Nations. Statistical versus Theoretical Relations in Economic Macrodynamics.

A memorandum prepared by Prof. Frisch for the Business Cycle Conference at Cambridge, England, July 18th - 20th, 1938, to discuss Professor J. Tinbergen's publications of 1938, for the League of Nations.

- - - - - - -

Introduction.

The present memorandum has been written rather hurriedly, and the text is therefore not as carefully polished as it ought to be in a manuscript ready for publication. It should, however, be clear enough to bring out my point of view.

The present memorandum does not discuss details of the various equations which Tinbergen has obtained and whose coefficients he has determined statistically. My main concern has been to discuss what equations of this type really mean, and to what extent they can be looked upon as "A Statistical Test of Business Cycle Theories". (The title of one of the volumes which Tinbergen has presented for discussion).

My conclusion is that the work which Tinbergen is now presenting is of paramount importance, perhaps the most important single step forward in Business Cycle Analysis of recent years. But I do not think that it can be looked upon as "A Test of Business Cycle Theories". The question of what connection there is between the relations we work with in theory and those we get by fitting curves to actual statistical data is a very delicate one. I think it has never been exhaustively and satisfactorily discussed. Tinbergen in his work hardly mentions it. He more or less takes it for granted that the relations he has found are in their nature the same as those of theory. See for instance his discussion in Vol. II P. 109 - 123 where he constantly refers to the coefficients of his equations and takes the signs and magnitudes of these as tests of whether certain theoretical contentions are right or wrong. This is, in my opinion, unsatisfactory. In a work of this sort, the connection between statistical and theoretical relations must be thoroughly understood and the nature of the information which the statistical relations furnish - although they are not identical with the theoretical relations - should be clearly brought out.

The present memorandum is an attempt to bring some contribution to this question. It will be divided into 7 sections viz.

1. Some remarks on terminology.
2. Functional equations and their solutions.
3. The irreducibility of a functional equation with respect to a set of functions.
4. Coflux and superflux relations. The nature of passive observations.
5. The autonomy of a functional equation. The nature of explanation, experimentation and reform.
6. Aberrations versus stimuli. Confluence analysis and shock-theory.
7. Interpretation of Professor Tinbergen's results.

1. <u>SOME REMARKS ON TERMINOLOGY.</u>

In any macrodynamic analysis there will be some constants or functions of time that are taken as <u>data</u> while others are considered as the <u>variates</u> to be "explained". A <u>determinate theory</u> is one that considers just as many independent relations as there are variates to be explained.

We shall use the expression "nature" or "constitution" of the system of phenomena studied as the whole of all those characteristics that describe the "way of functioning" of these phenomena. When we speak of the "structure" of the system, we think more specifically of those features of the "constitution" that can be quantitatively described. We speak for instance of the "structural equations" for the system. We do not intend however to draw any sharp line of demarcation between constitution and structure. The difference between them is only one of degree and one that is not very important. The precise definition of a structure is a matter of theorising although of course the leading ideas of the theoretical definitions will frequently be suggested by facts.

A disturbance is a deviation from that situation which should have existed as a consequence of the structure. In other words, it is something incompatible with the structure, something new and spontaneous introduced in addition to the structure. Such disturbances may be of two sorts: aberrations and stimuli. A stimulus is a disturbance that carries on its effects to the subsequent states of the system, - through the structural equations. In other words at any given moment it is the magnitudes of the variates including the stimuli that are taken as influencing the further evolution, that is, the stimuli act as a sort of permanently changing initial conditions. An aberration is also a departure from the value which a variate should have had according to the structure, but this departure acts only at the actual moment at which it occurs, it is a sort of instantaneous addition - unexplained by the structure - and without any consequence for the subsequent states. In other words it is the magnitudes of the variates exclusive of the aberrations that act as initial conditions for the subsequent states.

The existence of aberrations leads to the application of the methods of Confluence analysis.* The existence of stimuli leads to the shock-theory.**
There may also be mixed cases but I shall not go into this question here.

2. FUNCTIONAL EQUATIONS AND THEIR SOLUTION.

The structure of a macrodynamic system will be described by means of a number of functional equations. We shall in particular consider linear lag-equations, and taking as our variates the deviations from certain trend values,

---

* See the publication "Confluence Analysis" of the University Institute of Economics, Oslo.
** Publication in preparation at the Oslo Institute.

it will be sufficient for our purpose to consider homogeneous equations. Let $\bar{\bar{x}}_1(t) \ldots x_N(t)$ be a number of variates whose movement is to be explained, t designating time. Let $\bar{x}_1(t) \ldots \bar{x}_N(t)$ be trend values determined in some way or other, and let $x_1(t) \ldots x_N(t)$ be deviations from trend, i.e.

(2.1) $$x_i(t) = \bar{\bar{x}}_i(t) - \bar{x}_i(t)$$

Between the variates x_i we assume a number of relations (structure equations) of the form

(2.2) $$\Sigma_{i\theta}\alpha_{ki\theta}x_i(t - \theta) = 0 \qquad (k = 1,2\ldots)$$

k represents different equations, while the summation $i\theta$ represents the terms of each equation, i runs through all or some of the variate numbers 1 ... N, and θ runs through a certain range of lag numbers, in general different for each variate. The $i\,\theta$ range in each equation determines the nature of the terms involved, we shall call it the <u>form</u> of the equation, the α's are the <u>coefficients</u> of the equation. The distinction between the <u>form</u> and the <u>coefficients</u> of the equation is essential for the discussion in Section 4. A similar distinction may of course be made for more general types of functional equations.

For the discussion of the following Sections it is necessary to summarise some of the classical facts of the theory of linear lag-equations (difference equations).

A certain number of equations of the form (2.2) - equal or unequal to the number of variates N - are said to be linearly independent if it is impossible to deduce any one of the equations from the others no matter what the time shapes of the variates are. A necessary and sufficient condition for the independence of a set of m equations of the form (2.2), is that there should <u>not</u> exist any set of numbers $\lambda_1\lambda_2 \ldots \lambda_m$ not all zero, such that

(2.3) $$\Sigma_k\lambda_k\alpha_{ki\theta} = 0 \quad \text{for any } i\,\theta$$

In terms of the coefficients α the criterion can be formulated by considering the m rowed and M columned matrix

(2.4) $$\| \alpha_{ki\theta} \|$$

where all $i\,\theta$ combinations are written as columns, M being the number of different $i\theta$ combinations that exist in all the m equations and k representing the rows.

The equations are independent when and only when this matrix is of rank m. Or again the criterion can be formulated in terms of the moments $(\alpha_h\alpha_k)$ the summation being extended over all 10 combinations. The equations are independent when and only when the symmetric determinant of the magnitudes $(\alpha_h \alpha_k)$ is different from zero.

For each equation of the kind (2.2) that we add, we make less general the class of functions that satisfy the equations. If the number of equations becomes equal to N the system is <u>determinate</u>. This means that the nature of the solution has been restricted as much as it is possible to do so by means of functional time equations. It does not mean that the set of functions $x_1(t) \ldots x_N(t)$ is completely determined, a considerable amount of freedom is still left and will have to be determined by a set of <u>initial</u> conditions. But this determination is in point of principle different from that achieved by the functional equations. This is shown clearly by the fact that we cannot, say, replace the initial conditions by one or more <u>additional</u> equations of the form (2.2). Indeed if there are <u>more</u> than N independent equations of the form (2.2), there will in general exist no functions satisfying the system.

The <u>solution</u> of a determinate system of the form (2.2) can be achieved either directly or by means of expansions in series of complex exponentials. The direct method is applicable only in the simplest cases. Take as an example the system

$$\alpha_1 x_1(t) - \alpha_2 x_2(t) = 0 \tag{2.5}$$
$$\alpha_1 x_1(t) + \alpha_2 x_2(t - \theta) = 0$$

From this follows immediately $\alpha_1 x_1 + \alpha_1 x_1(t - \theta) = 0$

(2.6) hence $$x_1(t) + x_1(t - \theta) = 0$$

A solution of (2.6) is obtained by choosing arbitrarily the shape of x, over an interval of length θ and then repeating this shape antiperiodically for each subsequent θ interval. (2.6) shows that no more general form than this can be an X solution of (2.5). If we further put $x_2(t) = \frac{\alpha}{\alpha_1} x_1(t)$ we get a complete solution of (2.5). Obviously this is the most general form of the solution. Any function that can be a solution, must be a special case of this. The arbitrary shape of X_1 over the original θ interval is here the initial condition. Putting this equal to a sine function with period 2θ, we get, both for X_1 and X_2, over the complete t range, sine functions with this period.

In the more complicated cases one must resort to the indirect method. It consists in trying to satisfy the equations by expansions of the form

(2.7) $$x_i(t) = \Sigma_\gamma c_{i\gamma} e^{\gamma t} \qquad (i = 1,2,\ldots N)$$

where $c_{i\gamma}$ are constants and the summation over γ runs over certain values to be determined. Complex numbers are admitted both as c's and γ's.

Inserting (2.7) in (2.2) we get - if the system is determinate -

(2.8) $$\Sigma_{\gamma i\theta} \alpha_{ki\theta} c_{i\gamma} e^{\gamma(t-\theta)} = 0 \qquad (k = 1,2\ldots N)$$

Any number of exponential functions with different exponentials are linearly independent, therefore if the γ's are different, (2.8) cannot vanish identically in t unless the terms of (2.8) vanish separately for each γ. Assuming for the moment all the γ's to be different we see that we must have -

(2.9) $$\Sigma_{i\theta} \alpha_{ki\theta} c_{i\theta} e^{-\gamma\theta} = 0 \quad \text{for all } \gamma \text{ and } k.$$

For any given value of γ this is a system of linear homogeneous equations (k=1,2...N) in the set of N numbers If this system is to have a solution apart from the trivial $c_{1\gamma} = \ldots = c_{N\gamma} = 0$, the determinant of the coefficients must vanish, i.e. we must have

(2.10) $$/\Sigma_\theta \alpha_{ki\theta} e^{-\gamma\theta}/ = 0 \quad \text{for any } \gamma$$

k and i designate rows and columns respectively in the N rowed determinant (2.10). (2.10) is the characteristic equation whose roots $\gamma_0 \gamma_1, \ldots$ (in general infinite in number) give the exponents of the expression (2.7). Under very general conditions this expansion is valid even though some of the γ's are equal, the only difference being that in this case the multiple terms are replaced by a polynomial in t (of the order equal to $\kappa - 1$ if κ is the multiplicity of the γ- root) multiplied by the exponential in question. We need not consider this case here. The characteristic equation could also have been obtained by eliminating - in a way similar to that used to obtain (2.6) - a certain number of the variates in order to get a "final equation" in one or a few variates, and then forming the characteristic equation for this. This procedure is often useful when it is wanted to give a concrete interpretation of the mechanism of the solution, but in point of principle it is just as easy to form the characteristic equation directly as in (2.10).

It will be noted that the set of exponents as determined by (2.10) is the same for all the variates $x_1 \dots x_N$. In other words all the variates contain the same sort of components (if γ is a real number the component in question is a real exponential, if γ is a complex number its conjugate must also be a solution of the characteristic equation and these two terms together will form a real, damped, undamped or antidamped sine function). But the intensities with which the components occur in the various variates will be different. These intensities, - amplitudes - are represented by the numbers $c_{i\gamma}$. The distribution of these numbers and in particular the extent to which it is determined by the functional equation is essential for the interpretation of the relation between statistical and theoretical relations in economic macrodynamics.

For any given γ the corresponding numbers $c_{1\gamma} \dots c_{N\gamma}$ will - if (2.10) is of rank N - 1 for this value of γ - be uniquely determined apart from a common factor of proportionality c_γ. Indeed, the numbers $c_{1\gamma} \dots c_{N\gamma}$ will, when (2.10) is of rank N - 1, be proportional to the elements in a row of its adjoint (the elements of all these rows are proportional). There exists at least one row which does not consist exclusively of zeros and hence determined the proportions in question. Let $\hat{c}_{1\gamma} \dots \hat{c}_{N\gamma}$ be one such set of proportionality numbers. We may then put

(2.11)
$$c_{i\gamma} = c_\gamma \hat{c}_{i\gamma} \qquad \text{where } c_\gamma \text{ is an arbitrary number.}$$

This applies to any root γ that makes the rank of (2.10) N -1. Suppose that only such roots exist (the other cases do not alter those features of the amplitude distribution in which we are here interested).

Inserting (2.11) in (2.7.) we see that we can draw the following conclusions:

Each variate of the set of functions that is a solution of (2.2) can be expanded as a sum of trigonometric components. The frequencies and damping exponents of the components are determined by the equations (2.2), and so are the relative amplitudes, that is the ratio of the amplitude of a given component in one of the variates to that of the same component in another variate. But the absolute amplitudes are not determined by the equations (2.2). If these equations only are given, we may choose the absolute strengths of the various components in one of the variates arbitrarily (the choice of the numbers c_γ), but then the absolute strengths of these components in the other variates follow since their relations to the amplitudes of the components in the one variate we selected are determined by the equations (2.2). Briefly, the relative amplitude distribution is determined by the equations (2.2), but the absolute amplitude distribution has to be fixed by the initial conditions.

A similar situation exists for the phase distributions. Indeed the timing of a given component in one variate as compared with that of the same component in another variate is determined by the equations (2.2), but the timing of the various components in one selected variate must be fixed by the initial conditions.

By elimination processes similar to that used in obtaining a final equation, many new systems of equations may be deduced from (2.2). If the correspondence between the two systems is unique in the sense that the new system may be derived from the old and vice versa, with identical variates involved, the solutions of the two systems must be identical. In particular it is of interest to consider linear-elimination processes, that is processes where the form of the equations (the specification of the functions and lag-numbers that occur in the equation) is the same but the coefficients are changed. Any transformation of the form

$$\alpha^{*}_{ki\theta} = \Sigma_h \, \psi_{kh} \, \alpha_{hi\theta} \tag{2.12}$$

where ψ_{kh} is a non-singular matrix, independent of $i\theta$, will furnish a new system of equations, that are independent if the old system is, and has exactly the same set of functions as its solution, - and vice versa.

We shall now discuss in somewhat greater detail these various equations that have the same solutions, and introduce a classification of them which is important for our purpose. In particular we shall consider equations which have the same form but different coefficients.

3. THE IRREDUCIBILITY OF A FUNCTIONAL EQUATION WITH RESPECT TO A SET OF FUNCTIONS.

When we compare a functional equation involving one or several functions with particular set of functions, there are two questions to be asked: Does the set of functions satisfy the equation and does it satisfy this equation only?

Obviously, the set will satisfy all equations which - involving identical functions - can be derived from the first equation, so we are only interested in knowing whether the set satisfies some other equation which is independent of the first. Furthermore we shall not consider all other conceivable equations but only those which are of the same form as the first but have different coefficients. In the case of homogeneous equations of the form (2.2) this means that for any given one of these equations (any given k) we are interested in knowing whether a particular set of functions considered satisfies not only this equation but also another with the same i θ range but with coefficients that are non-proportional to those of the first equation. If this is so, we shall say that the first equation is reducible with respect to this set of functions, if not it is irreducible. Thus an irreducible equation of the form (2.2) is one whose coefficients are uniquely determined and allow of no degree of freedom if the equation is to be satisfied by this set of functions (apart from the arbitrary factor of proportionality which is always present in the case of a homogeneous equation). It is clear that the property of irreducibility must be important when we are studying the nature of those equations that can be determined from the knowledge of the time shapes of the functions that are to satisfy the equations.

Obviously the first equation in the above definition is reducible , the second is also reducible. The set of functions involved in the definition may be specified in great detail or only very broadly as a general class of functions.

A similar definition may be established for a system of equations but we shall only need it for a single equation.

If an equation is given, we may consider the class of all those sets of functions (satisfying the equation) which have the property that the equation is irreducible with respect to those sets of functions. This class we may call the irreducibility class of the equation.

Let us consider some simple propositions and some examples that will help us to visualize the nature of this irreducibility definition. In the first place it is easy to see that there cannot exist two or more equations of the same form which are both irreducible with respect to the same set of functions. But if the two equations are of different forms (e.g. with different lag-numbers) each of them may be irreducible for the same set of functions.

In the second place we notice that any functional equation is irreducible with respect to the most general set of functions that satisfy this equation. Indeed, if the set of functions should also satisfy another equation - independent of the first - this would represent a restriction to the set, so that it could not actually be the most general set that satisfied the first equation. But if we consider a set of functions that satisfy two independent equations, neither of the equations need be irreducible with respect to this set. This is indeed a more special set of functions and the requirement that this set shall be a solution is less rigorous and therefore places less restrictions on the coefficients of the equation.

As a more particular example let us consider the equation (2.6). Pure sine functions with period 2θ is a solution, and for functions of this sort the equation is irreducible, because there do not exist any values of p and q which will make the equation

$$px_1(t) + qx_1(t - \theta) = 0 \tag{3.1}$$

an equation satisfied by pure sine functions of period 2θ, except the values $p = q$. And in this case the equation is the same as (2.6).

On the other hand take the equation

$$0.6x_1(t) + x_1(t - \theta) + 0.4x_1(t - 2\theta) = 0 \tag{3.2}$$

This equation is also satisfied by a sine function of period 2θ (which is easily seen by insertion), but it is not irreducible with respect to this function. The equation would also be satisfied by this function if we let the first coefficient be 0.9 and the last 0.1, or quite generally if the sum of the first and last coefficients are equal to the middle coefficient. In this case the coefficients of the equation have a one dimensional degree of arbitrariness (even apart from the arbitrary factor of proportionality which is always present in the homogeneous equations).

The following is a general rule about the reducibility of equations of the form (2.2).

(3.3). Rule about reducibility: If the functions with respect to which reducibility is defined are made up of n exponential components (two complex exponentials correspond to one damped, undamped or antidamped sine function), the equation is certainly reducible - and hence its coefficients are affected in a more or less arbitrary manner - if it contains more than n+1 terms. And it may be reducible even if it contains n+1 terms or less.

Let us first consider as an example the following three term equation in one function

$$\alpha_1 . x(t - \theta_1) + \alpha_2 . x(t - \theta_2) + \alpha_3 (t - \theta_3) = 0 \tag{3.4}$$

If x is simply an exponential $x(t) = Ce^{\gamma t}$, the left-hand side of (3.4) becomes $Ce^{\gamma t}[\alpha_1 e^{-\gamma\theta_1} + \alpha_2 e^{-\gamma\theta_2} + \alpha_3 e^{\gamma\theta_3}]$
In order that this expression should vanish identically in t it is necessary and sufficient that the bracket should disappear. This leaves a one dimensional arbitrariness in the α's even apart from their arbitrary common factor of proportionality.

If the function considered is of the form

$$x(t) = Ae^{\beta t} \sin(a + \alpha t) \tag{3.5}$$

it is equivalent to two exponential components, and since the number of terms in the equation is only 3, the equation may be irreducible. But it may also be reducible if the lag-numbers satisfy certain special conditions. Inserting from (3.5) into the left-hand side of (3.4) we get

$$Ae^{\beta t}\sin(a + \alpha t)[\alpha_1 C_1 + \alpha_2 C_2 + \alpha_3 C_3] + Ae^{\beta t}\cos(a + \alpha t)[\alpha_1 \tau_1 + \alpha_2 \tau_2 + \alpha_3 \tau_3] \tag{3.6}$$

where

$$C_i = e^{-\beta\theta_i} \cos \alpha \theta i \qquad \tau_i = e^{-\beta\theta_i} \sin \alpha \theta i \tag{3.7}$$

In order that (3.6) shall vanish identically in t it is necessary and sufficient that the two brackets should disappear separately (because the two time functions in front of the brackets are linearly independent.) I.e. we must have

$$\begin{aligned} \alpha_1 C_1 + \alpha_2 C_2 + \alpha_3 C_3 &= 0 \\ \alpha_1 \tau_1 + \alpha_2 \tau_2 + \alpha_3 \tau_3 &= 0 \end{aligned} \tag{3.8}$$

These are two equations in the α's. If the coefficients of the two equations are proportional, i.e. if

(3.9) $$\frac{c_1}{\tau_1} = \frac{c_2}{\tau_2} = \frac{c_3}{\tau_3}$$

a set of α's that satisfies one of the equations would automatically satisfy the other. Hence there would again be only one condition for the three α's consequently the α's would have a one dimensional arbitrariness, even apart from the usual factor of proportionality. The condition (3.9) is equivalent to the condition that all the three two-rowed determinants in the matrix of coefficients in (3.8) should vanish (if any two of these determinants vanish, the third vanishes automatically).

Since

(3.10) $$\begin{vmatrix} c_i & c_j \\ \tau_i & \tau_j \end{vmatrix} = e^{-\beta(\theta_i + \theta_j)} \sin \alpha(\theta_j - \theta_i)$$

we see that in terms of the lag-numbers the condition (3.9) is reduced to

(3.11) $$\theta_2 - \theta_1 = \frac{h\pi}{\alpha} \qquad \theta_3 - \theta_1 = \frac{k\pi}{\alpha}$$

where h and k are integers ($h = k$ or $h = 0$ or $k = 0$ represent trivial cases). Thus, if (3.11) is fulfilled, (3.4) is reducible with respect to (3.5).

A general criterion for the case when even the $(n + 1)$ terms equation is reducible with respect to a series consisting of n exponential components, is provided by the n rowed and $(n + 1)$ columned matrix

(3.12) $$\left\| c_{i\gamma} e^{-\gamma\theta_{i_1}}, \; c_{i\gamma} e^{-\gamma\theta_{i_2}} \ldots c_{j\gamma} e^{-\gamma\theta_{j_1}}, \; c_{j\gamma} e^{-\gamma\theta_{j_2}} \right\|$$

where $i, j \ldots$ are the affixes of the variables $x_i, x_j \ldots$ which occur in the equation considered, and $\theta_{i_1}, \theta_{i_2} \ldots \theta_{j_1}, \theta_{j_2} \ldots$ are the lag-numbers. The rows of (3.12) are produced by letting γ run through the characteristic numbers of which we here suppose that there exist n. For the $(n + 1)$ term equation in question to be reducible with respect to the set of functions considered, it is necessary and sufficient that the matrix (3.12) should be less than n. More precisely: if it is of rank $r \leqq n$ (which is a criterion that depends only on the nature of the functions in question and the distribution of the lag-numbers) the equation will have an $(n - r)$ dimensional reducibility, i.e. its coefficients will have an $(n - r)$ dimensional degree of arbitrariness, in addition to the arbitrary factor of proportionality associated with the homogeneity of the equation.

4. COFLUX AND SUPERFLUX EQUATIONS. THE NATURE OF PASSIVE OBSERVATIONS.

If a determinate system of the form (2.2) is given, it is of particular interest to consider the reducibility of the various equations with respect to that class of functions which is a solution of the complete system. This, of course, is a much more special class of functions than that which satisfies each equation taken separately, and the reducibility of the equation is correspondingly higher. The specialisation of the functions is still further increased by the initial conditions. We have indeed seen that even though the solution of the equations themselves may contain a large - perhaps infinite - number of components, the equations do not say anything about the absolute amplitude distribution. It may indeed happen that in the actual solution all components will disappear except, say, one which is a pure sine curve. In this case all the original equations that consisted of more than three terms would certainly be reducible, and even some of the three-term equations might be reducible. An equation which is irreducible with respect to the set of functions that forms the actual solution of the complete system (including those determined by the initial conditions) we shall call a coflux equation. The others - those that are reducible with respect to their set of functions - will be called superflux equations. These latter equations are of course in a particular sense irreducible, but with respect to more general classes of functions. If any one of them is not reducible for any more special class of function it is at least irreducible for that class which consists of its own most general solution. The word "flux" in this connection suggests that the reducibility is here defined with respect to the time shape - the "flux" - actually possessed by the phenomena.

The notion of coflux relations is fundamental when we ask what sorts of equations it is possible to determine from the knowledge of the time shapes that are actually produced. The answer is obviously that all coflux equations and no other equations are discoverable from the knowledge of the time shapes of the functions that form the actual solution.

Indeed all other equations will have coefficients with at least a one-dimensional degree of arbitrariness. If an attempt were made to fit such an equation to the data, the coefficients would be of the $\frac{0}{0}$ form when no errors (aberrations) were present, and otherwise they would have a fictitious determinateness, their magnitudes being determined solely by the errors, and not by the structure.

This is the nature of passive observations, where the investigator is restricted to observing what happens when all equations in a large determinate system are actually fulfilled simultaneously. The very fact that these equations are fulfilled prevents the observer from being able to discover them, unless they happen to be coflux equations, that is, irreducible with respect to the functions that form the actual solution.

But why bother about these other equations that are not discoverable through passive observations?

The answer is that some of these other equations frequently have a higher degree of "autonomy" than the coflux equations, and are therefore very well worth knowing. The "autonomy" of an equation is not, like the irreducibility a mathematical property of a closed system like (2.2), but is built on some sort of knowledge outside this system. I shall now proceed to a discussion of this point.

5. THE AUTONOMY OF A FUNCTIONAL EQUATION. NATURE OF "EXPLANATIONS", EXPERIMENTATION AND REFORM.

Suppose that, from a knowledge of the time shapes of the two functions $x_1(t)$ and $x_2(t)$ I have determined a relation of the form

$$(5.1) \qquad x_1(t) = ax_1(t - \theta) + bx_2(t) + cx_2(t - \theta_2)$$

What does this equation mean? It means that so long as x_1 and x_2 continue to move with the same time shapes as they have had in the past I can compute the value of x_1 at any point of time t from the knowledge of x_2 at this same point and x_1 and x_2 at certain earlier moments as indicated in the formula. In other words the equation is simply a description of the "routine of change" which x_1 and x_2 follow. The equation determined in this empirical way does not state that if a situation occurs where $x_1(t-\theta_1)$, $x_2(t)$ and $x_2(t-\theta_2)$ have some arbitrary values. I can again compute $x_1(t)$ by (5.1) To assume that (5.1) should hold good for any values whatsoever inserted for the variables on the right-hand side of the equation would indeed imply that I conceived of the possibility of another structure than the one which prevailed when the equation (5.1) was determined. For instance, if the original structure was taken as defined by two equations of the form (2.2), I could not conceive of a free variation of the variates on the right-hand side of (5.1) without giving up at least one of the two structural equations that determine the course of x_1 and x_2. But that would mean giving up the very assumption on which (5.1) was determined.

This situation can also be interpreted in terms of irreducibility. If I conceive of the possibility that the constants, a,b, and c in (5.1) may have definite values, I must also conceive of the existence of some time shapes of x_1 and x_2 for which (5.1) becomes an irreducible equation, that is, has determinate coefficients without any arbitrariness. And the same applies to any other structural equation.

In a big system of structural equations it would be quite exceptional if all the equations should be irreducible with respect to that particular solution which turns out to be the final one. We only have to think of a case where the initial conditions are such that only one single component is left with an amplitude different from zero, while many of the structural equations contain a large number of terms. The fact that I reckon with such a system of equations, must mean that I conceive of the possibility that the structure may have been different from what it actually is, thus giving a chance of producing a time shapes complicated enough to make the big structural equations irreducible with respect to these time shapes.

But when we start speaking of the possibility of a structure different from what it actually is, we have introduced a fundamentally new idea. The big question will now be: in what directions should we conceive of a possibility of changing the structure. There is nothing in the nature of the equations that describe the actual structure, which can suggest an answer. It is true that if a system of equations is given, it would be natural to imagine in turn all equations omitted except one, this remaining equation would then certainly be irreducible with respect to the general class of functions which now satisfy the equation (see the second example in Section 2). But this solution is only apparent, because there exist an infinity of ways of writing the system of structural equations. (Compare for instance the transformation (2.12))

To get a real answer we must introduce some fundamentally new information. We do this by investigating what features of our structure are in fact the most autonomous in the sense that they could be maintained unaltered while other features of the structure were changed. This investigation must use not only empirical but also abstract methods. So we are led to constructing a sort of super-structure, which helps us to pick out those particular equations in the main structure to which we can attribute a high degree of autonomy in the above sense. The higher this degree of autonomy, the more fundamental is the equation, the deeper is the insight which it gives us into the way in which the system functions, in short, the nearer it comes to being a real explanation. Such relations form the essence of "theory".

Once such a basic system of structural equations to which we can attach the label "autonomous" has been selected, it is easy to derive others that have a greater or lesser degree of autonomy. Equations that are obtained by long elimination processes, based on several autonomous equations will have a low degree of autonomy, they will in fact depend on the preservation of a great many features of the total system.

The coflux relations that can be determined by observation of the actual time shapes may or may not come near to resembling an autonomous relation, that depends on the general constitution of the phenomena studied. To give two extreme examples: the demand function for a consumers commodity as depending on price and income and perhaps on some secondary variables will, if the coefficients can be determined with any degree of accuracy, come fairly near to being an autonomous relation. It will not be much changed by a change in monetary policy, in the organisation of production etc. But the time relation between the Harvard A, B, and C curves is a pure coflux relation, with only a small degree of autonomy.

Such I believe is in essence the relation between the equations of pure theory and those that can be determined by passive observations.

If the situation is such that the coflux relations are far from giving information about the autonomous structural relations, recourse must be had to experimentation, that is one must try to change the conditions so that one or more of the structural equations is modified. In economics the interview method is a substitute - sometimes bad, sometimes good - for experimentation.

If the results of the investigation are to be applied for economic political purposes - for reforming the existing economic organisation - it is obviously the autonomous structural relations we are interested in.

6. ABERRATIONS VERSUS STIMULI. CONFLUENCE ANALYSIS AND SHOCK-THEORY.

The existence of aberrations does not necessarily involve any important consequences for the theoretical analysis, it only concerns the statistical technique, but in this respect it is important. The existence of stimuli entails much more far-reaching consequences. The total time shape will now be more or less transformed, for instance damped cycles will become undamped in the long run, but will have a disturbing effect over shorter intervals. The timing between the cycles may be changed from what it is in the stimulus-free system, and entirely new cycles, pure cumulation cycles will emerge. These consequences cannot be discussed in detail here.

7. INTERPRETATION OF PROFESSOR TINBERGEN'S RESULTS.

All the way through his work Tinbergen uses approximations by which the time equations are reduced to linear forms. This is certainly admissible in a first approximation but the consequences should be clearly recognised. If the linear approximations are used for as many equations as are needed to make the system determinate (which is what Tinbergen aims at doing: "...we must continue this procedure until the number of relations obtained equals the number of phenomena..." "Business Cycles...." p.7) - only those features of the time series are taken account of that can be approximated to by fitting to the data that type of solution which a linear system of equations admits of, namely a number of trigonometric components (exponentials or damped, undamped or antidamped sine functions or as exceptional cases such functions multiplied by polynomials) of the time series over the interval considered. In itself there is nothing objectionable in this but it means that the significance of the results must be interpreted in the light of the various algebraic facts of the preceding sections. These become relevant with the same approximation as that involved in Tinbergen's calculations.

This being so it is clear that it is only coflux relations that are determined by Tinbergen, and the lack of agreement between these equations and those of pure theory cannot be taken as a refutation of the latter. Any number of examples could be given of statements that are in need of very much qualification on this ground. A case in point is that discussed on page 111 in "Business Cycles" or perhaps even better the attempt on page 26 to get an equation for consumers outlay. The only result of the various attempts made here is to shift from one to another amongst an infinite number of coflux equations. By a suitable choice of the variate and lag-numbers introduced one can produce practically any coefficients one likes. A computation from series made upfor p of a small number of trigonometric components shows this immediately. The reasons for discarding some of the equations (p26) are quite unsatisfactory. No other reasons seem to be given than the fact that the coefficients do not work out as the author likes. In my opinion all these equations are acceptable when interpreted as what they really are: a number of coflux equations. But none of them can, I believe, be taken as an expression of the autonomous structural equation that will characterize demand.

In concluding this memorandum, I want to stress again what I mentioned in the introduction, namely, the importance of the results obtained by Tinbergen. They will have to be taken as starting point for any further investigation aiming at obtaining limits or other sorts of information concerning the structural coefficients.

17.7.38.

(signed) RAGNAR FRISCH.

2. League of Nations. Mr. Tinbergen's reply to Professor Frisch's note on "Statistical Versus Theoretical Relations in Economic Macrodynamics"

I. Summary of Professor Frisch's note.

In sections 1 and 2 of his note Prof. Frisch gives some definitions concerning the logical structure of business cycle theories which are so clear that they need no further comment. As regards what follows, I should like to draw special attention to the important distinction he makes between two sorts of disturbances, viz. aberrations and stimuli. As regards the former it is supposed that they do not exert any influence on the further course of events, whereas the latter are assumed to have an influence.

Sections 3 and 4 introduce the notions of reducible and irreducible equations, (with respect to a set of functions) and coflux and superflux relations. This could be most easily represented to the non-mathematical reader of my two reports as a consequence of multicollinearity. It has been remarked in these reports that the determination of regression coefficients becomes impossible if some of the "explanatory" variables are linearly connected (the simplest case being proportionality between two of these variables). Frisch's remarks could be interpreted by saying that there is a systematic tendency to such linear connections as soon as business cycle research is entered upon. These linear connections are the consequence of the other equations of the system. Putting in a familiar but oversimplified form, one could say that "all business cycle curves are more or less sines or waves, and that therefore the danger of multicollinearity is permanently present."

Frisch first defines "reducibility of an equation with respect to a set of functions" irrespective of the business cycle mechanism, and only introduces the latter when speaking of coflux and superflux relations. An equation

$$p = a\,x + b\,y + c\,z \qquad (1)$$

a reducible with respect to a certain set of (time) functions p, x, y, and z, if there is a further linear relation between p, x, y and z, say

$$Ax + By + Cz = 0 \qquad (2)$$

by which it is possible to replace (1) by an infinite number of other relations, e.g.,

$$p = (a - 3A)\,x + (b - 3B)\,y + (c - 3C)\,z \qquad (3)$$

In these formulae it may e.g. also happen that y is nothing but a lagged value of x, say $y = x_{-1}$. Frisch's definition is more precise in some respects but for our argument this is not so important.

On the other hand, (1) is said to be irreducible with respect to p,x,y and z if a relation of the type (2) (possibly including p) does not exist. Then (1) cannot be replaced by another equation which is independent of (1),as is (3).[x]

Irreducible equations with respect to such functions p,x,y,z which form the solution of the whole system of business cycle equations are called coflux relations. Reducible ones are called superflux relations. In this case (i.e. when we are speaking of the solutions of the whole system of equations) there are as many equations as there are variables, but they may of course show lags which are different from those occurring in (1).

About the functions now under consideration, Frisch remarks that they form, "of course,..... a much more special class of functions than that which satisfies each equation taken separately, and the reducibility of the equations (is) correspondingly higher."

Prof. Frisch observes further that only irreducible (i.e. coflux) relations can be determined fairly exactly, since the reducible ones are, from the regression analysis viewpoint, indeterminate.

In section 5 the notion of autonomy of a relation is brought in; it is,as far as I can see, the same as what I mean by direct relations (in contradiction to indirect relations which are obtained by one or more elimination steps). This notion is essentially an economic one. In order to "explain", or to study the consequences of policy, one has to know these autonomous relations. Of these relations Frisch already says on page 28 (section 4): autonomous relations are often superflux relations (i.e. reducible ones with respect to the solutions of the system).[xx] As extreme

[x] An example of a dependent equation would be $4p = 4\,ax + 4\,by + 4\,cz$.

[xx] He says it in the form: "....frequently, some of the other equations (i.e., the reducible ones) have a higher degree of autonomy."

examples of autonomous and non-autonomous relations Prof. Frisch gives the following two cases: the demand function for a consumers' commodity as depending on price and income and perhaps on some secondary variables will, if the coefficients can be determined with any degree of accuracy, come fairly near to being an autonomous relation. It will not be much changed by a change in monetary policy, in the organisation of production, etc. But the time relation between the Harward A, B and C curves is a pure coflux relation, with only a small degree of autonomy.

After some short remarks, in section 6, on aberrations versus stimuli, Prof. Frisch gives his conclusions about the League work in section 7. The following sentences may especially be reproduced:

> "All the way through his work Tinbergen uses.... linear forms..... This being so it is clear that it is only coflux relations that are determined.....; and the lack of agreement between these equations and those of pure theory cannot be taken as a refutation of the latter.
>
> A case in point (viz. where very much qualification is needed J. T.) is..... the attempt on p. 26, Business Cycles, U.S.A., to get an equation for consumers' outlay. The only result of the various attempts made here is to shift from one to another amongst an infinite number of coflux equations. By a suitable choice of the variates and lag numbers introduced, one can produce practically any coefficient one likes..... The reasons for discarding some of the equations (p.26) are quite unsatisfactory; no other reasons seem to be given than the fact that the coefficients do not work out as the author likes.
>
> <u>In my opinion all these equations are acceptable</u> when interpreted as what they really are: a number of coflux equations. But none of them can, I believe, be taken as an expression of <u>the</u> autonomous structure equation that will characterise demand."

II. <u>Questions to Professor Frisch.</u>

(a) Must not the degree of autonomy of a relation be determined by economic considerations?

I made this the basis of my work by starting always from a relation the variables of which were based upon a priori economic considerations. I consider this as a guarantee that if any result at all is obtained this is an autonomous relation (or a direct one, as I would call it). Would you advise a different method?

(b) Why is it probable that autonomous relations are often reducible?

(c) Is not the procedure used in establishing the equation for consumers' outlay very much the same as that used for establishing demand functions for single commodities?

(d) Which equations in the system would you accept as autonomous ones and which not?

III. Reasons why I feel surer about our relation than Prof. Frisch does.

(a) As has already been expressed to some extent in question II (a) above, I think that most, if not all, of my relations are autonomous, or almost so, because I started each section with a priori economic considerations. Professor Frisch does not say on what grounds a relation like (1) above should be obtained, but in my opinion it makes all the difference whether a priori economic reasoning or, rather superficial observation "without theoretical prejudice" as in the Harward barometer case, is used.

(b) In addition I am less afraid than Prof. Frisch of the consequences of the other linear relations, like (2) above, since:

(1) many contain extraneous variables equivalent to stimuli (like Au, I_{Bp}, f,h);

(2) some very important relations are non-linear, viz. the n - equation (3.7) and, in principle at least, also the q - equation (3.5) and q_B - relation (3.6), where variable o stands for non-linear expression in other variables;*

(3) the freedom in the choice of lags and coefficients Prof. Frisch speaks of is considerably reduced once economic reasoning is accepted as a basis. Negative lags and

---

*As a consequence of points (1) and (2), interest rates and banking variables are all rather clearly "deformed" by Au, prices by o, profit income and production figures by stock exchange movements. This makes a discrimination between influences of these three groups easier than it would be in theory, since the movements of Au, o, and n are very different.

lags of more than some definite period, or coefficients of one sign are often prohibited.

Taking the example chosen by Prof. Frisch, viz., the determination of consumers' outlay, I think a correct impression of what is needed is not given when Prof. Frisch says: "no other reasons (for discarding some of the equations) seem to be given than the fact that the coefficients do not become what the author would like to see". What in fact the author likes is to get economically significant relations; and therefore he required that:

(i) the marginal propensity of workers should be larger than the one for non-workers;

(ii) the influence of Pareto's coefficient α if any, should be positive;

(iii) the influence of last year's income, if any, should be positive.

All this seems to me sound discrimination. The doubt which remains is recognised (viz. that the marginal propensity to consume for workers may be anything between zero and 0.10) and (this Prof. Frisch did not know) the influence of this uncertainty on the final equation is calculated and will be shown.

(c) Finally I introduced a number of <u>a priori</u> coefficients in cases where that seemed possible (e.g. the price equations (3.5) and (3.6), and the proportion between the first two coefficients in equation (3.7)).

(d) It may be added that a lack of agreement between our equations and "those of pure theory" is not alleged; merely a lack of agreement with those of some theories.

IV. <u>Proposals for changes in the text.</u>

Since Professor Frisch's note contains very valuable and important remarks of a systematic nature it seems worth while embodying large parts of it in the (enlarged) introduction. Apart from my own terminology Professor Frisch's should be mentioned. The reasons why I am less afraid of the dangers Prof. Frisch mentions than he is, should be stated; where uncertainties exist, their influence will be estimated (this has, in fact, already been done since the reports were printed; and the results were shown at the Cambridge meeting).

(Signed) J. Tinbergen.

3. "The Probability Approach in Econometrics" by Trygve Haavelmo. Cowles Commission Papers. New Series. No. 4. 1944. Section 8 of Chapter II.

The Autonomy of an Economic Relation.

Every research worker in the field of economics has, probably, had the following experience: When we try to apply relations established by economic theory to actually observed series for the variables involved, we frequently find that the theoretical relations are "unnecessarily complicated"; we can do well with fewer variables that assumed a priori. But we also know that, when we try to make predictions by such simplified relations for a new set of data, the relations often break down, i.e., there appears to be a break in the structure of the data. For the new set of data we might also find a simple relation, but a different one. Even if no such breaks appear, we are puzzled by this unexpected simplicity, because, from our theoretical considerations we have the feeling that economic life is capable of producing variations of a much more general type. Sometimes, of course, this situation may be explained directly by the fact that we have included in our theory factors which have no potential influence upon the variables to be explained. But more frequently, I think, the puzzle is a result of confusing two different kinds of variations of economic variables, namely hypothetical, free variations, and variations which are restricted by a system of simultaneous relations.

We see this difference best by considering the rational operations by which a theoretical system of relations is constructed. Such systems represent attempts to reconstruct, in a simplified way, the mechanisms which we think lie behind the phenomena we observe in the real world. In trying to rebuild these mechanisms we consider one relationship at a time.

Suppose, e.g., we are considering n theoretical variables $x_1', x_2', \ldots, x_n'$, to be compared with n observational variables $x_1, x_2, \ldots, x_n$, respectively. We impose certain relations between the n theoretical variables, of such a type that we think the theoretical variables, so restricted, will show some correspondence with the observed variables.

Let us consider one such particular relation, say $x_1' = f(x_2', \ldots, x_n')$. In constructing such a relation, we reason in the following way: If x_2' be such and such, x_3' such and such, etc., then this implies a certain value of x_1'. In this process we do not question whether these "ifs" can actually occur or not. When we impose more relations upon the variables, a great many of these "ifs" which were possible for the relation $x_1' = f$ separately, may be impossible, because they violate the other relations. After having imposed a whole system of relations, there may not be very much left of all the hypothetical variation with which we started out. At the same time, if we have made a lucky choice of theoretical relations, it may be that the possible variations that are left over agree well with those of the observed variables.

But why do we start out with much more general variations than those we finally need? For example, suppose that the Walrasian system of general-equilibrium relations were a true picture of reality; what would be gained by operating with this general system, as compared with the simple statement that each of the quantities involved is equal to a constant? The gain is this: In setting up the different general relations we conceive of a wider set of possibilities that might correspond to reality, were it ruled by one of the relations only. The simultaneous system of relations gives us an explanation of the fact that, out of this enormous set of possibilities, only one very particular one actually emerges. But once this is established, could we not then forget about the whole process, and keep to the much simpler picture that is the actual one? Here is where the problem of autonomy of an economic relation comes in. The meaning of this notion, and its importance, can, I think, be rather well illustrated by the following mechanical analogy:

If we should make a series of speed tests with an automobile, driving on a flat, dry road, we might be able to establish a very accurate functional relationship between the pressure on the gas throttle (or the distance of the gas pedal from the bottom of the car) and the corresponding maximum speed of the car. And the knowledge of this relationship might be sufficient to operate the car at a prescribed speed. But if a man did not know anything about automobiles, and he wanted to understand how they work, we should not advise him to spend time and effort in measuring a relationship like that. Why? Because (1) such a relation leaves the whole inner mechanism of a car in complete mystery, and (2) such a relation might break down at any time, as soon as there is some disorder or change in any working part of the car. (Compare this, e.g., with the well-known lag-relations between the Harvard A-B-C-curves.) We say that such a relation has very little autonomy.* because its existence depends upon the simultaneous fulfilment of a great many other relations, some of which are of a transitory nature. On the other hand, the general laws of thermodynamics, the dynamics of friction, etc., etc., are highly autonomous relations with respect to the automobile mechanism, because these relations describe the functioning of some parts of the mechanism irrespective of what happens in some other parts.

Let us turn from this analogy to the mechanisms of economic life. Economic theory builds on the assumption that individuals' decisions to produce and to consume can be described by certain fundamental behavioristic relations, and that, besides, there are certain technical and institutional restrictions upon the freedom of choice (such as technical production functions, legal restrictions, etc.).

A particular system of such relationships defines one particular theoretical structure of the economy, that is to say, it defines a theoretical set of possible simultaneous sets of value or sets of time series for the economic variables. It might be necessary - and that is the task of economic theory - to consider various alternatives to such systems of relationships, that is, various alternative structures that might, approximately, correspond to economic reality at any time. For the "real structure" might, and usually does, change in various respects.

To make this idea more precise, suppose that it be possible to define a class, Ω, of structures, such that one member or another of this class would, approximately, describe economic reality in any practically conceivable situation. And suppose that we define some non-negative measure of the "size" (or of the "importance" or "credibility") of any subclass, ω in Ω, including Ω itself, such that, if a subclass contains completely another subclass, the measure of the former is greater than, or at least equal to, that of the latter, and such that the measure of Ω is positive. Now consider a particular subclass (of Ω), containing all those - and only those - structures that satisfy a particular relation "A". Let ω_A be this particular subclass. (E.g., ω_A might be the subclass of all those structures that satisfy a particular demand function "A"). We then say that the relation "A" is autonomous with respect to the subclass of structures ω_A. And we say that "A" has a degree of autonomy which is the greater the larger be the

*This term, together with many ideas to the analysis in the present section, I have taken from a mimeographed paper by Ragnar Frisch: "Statistical versus Theoretical Relations in Economic Macro-Dynamics" (Mimeographed memorandum prepared for the Business Cycle Conference at Cambridge, England, July 18-20, 1938, to discuss J. Tinbergen's publication of 1938 for the League of Nations.)

"size" of ω_A as compared with that of Ω.

The principal task of economic theory is to establish such relations as might be expected to possess as high a degree of autonomy as possible.

Any relation that is derived by combining two or more relations within a system, we call a confluent relation. Such a confluent relation has, of course, usually a lower degree of autonomy (and never a higher one) than each of the relations from which it was derived, and all the more so the greater the number of different relations upon which it depends. From a system of relations, with a certain degree of autonomy, we may derive an infinity of systems of confluent relations. How can we actually distinguish between the "original" system and a derived system of confluent relations? That is not a problem of mathematical independence or the like, more generally, it is not a problem of pure logic, but a problem of actually knowing something about real phenomena, and of making realistic assumptions about them. In trying to establish relations with high degree of autonomy we take into consideration various changes in the economic structure which might upset our relations, we try to dig down to such relationships as actually might be expected to have a great degree of invariance with respect to certain changes in structure that are "reasonable".

It is obvious that the autonomy of a relation is a highly relative concept, in the sense that any system of hypothetical relations between real phenomena might itself be deducible from another, still more basic system, i.e., a system with still higher degree of autonomy with respect to structural changes.

The construction of systems of autonomous relations is, therefore, a matter of intuition and factual knowledge, it is an art.

What is the connection between the degree of autonomy of a relation and its observable degree of constancy or persistence?

If we should take constancy or persistence to mean simply invariance with respect to certain hypothetical changes in structure, then the degree of constancy and the degree of autonomy would simply be two different names for the same property of an economic relation. But if we consider the constancy of a relation as a property of the behavior of actual observations, then there is clearly a difference between the two properties, because then the degree of autonomy refers to a class of hypothetical variations in structure, for which the relation would be invariant, while its actual persistence depends upon what variations actually occur. On the other hand, if we always try to form such relations as are autonomous with respect to those changes that are in fact most likely to occur, and if we succeed in doing so, then, of course, there will be a very close connection between actual persistence and theoretical degree of autonomy. To bring out these ideas a little more clearly we shall consider a purely formal set-up.

Suppose we have an economic system, the mechanism of which might be characterized by the variations of n measurable quantities $x_1, x_2, \ldots, x_n$. Suppose that the structure of this mechanism could be described by a system of $m < n$ equations,

$$f_i(x_1, x_2, \ldots, x_n) = 0 \qquad (i = 1, 2, \ldots, m). \tag{8.1}$$

$(n - m)$ of the variables - let them be $x_{m+1}, x_{m+2}, \ldots, x_n$ - are assumed to be given from outside. From the system (8.1) it might, e.g., be possible to express each of the first m variables uniquely in terms of the n - m remaining ones. Let such a solution be

$$\begin{aligned} x_1 &= u_1(x_{m+1}, x_{m+2}, \ldots, x_n), \\ x_2 &= u_2(x_{m+1}, x_{m+2}, \ldots, x_n), \\ &\ldots\ldots\ldots\ldots \\ x_m &= u_m(x_{m+1}, x_{m+2}, \ldots, x_n). \end{aligned} \tag{8.2}$$

The system (8.2) would describe the covariations of the variables just as well as would the original system (8.1). But suppose now that there should be a change in structure of the following type: One of the functions f, in (8.1), say f_1, is replaced by another function, say f_1', while all the other relations in (8.1) remain unchanged. In general, this would change the whole system (8.2), and if we did not change the system (8.2) (e.g., because we did not know the original system (8.1)), some or all of its relations would show lack of constancy with respect to the observations that would result from the new structure. On the other hand, the last m - 1 equations in (8.1) would - by definition - still hold good, unaffected by the structural change. It might be that, as a matter of fact, one or two particular equations in (8.1) would break down very often, while the others remained valid. Then any system (8.2) corresponding to a fixed system (8.1) would show little persistence with respect to the actual observations.

In this scheme the variables $x_{m+1}, x_{m+2}, \ldots, x_n$, were, in point of principle, free: they might move in any arbitrary way. This includes also the possibility that, e.g., all these free variables might move as certain well-defined functions of time, e.g.,

$$(8.3) \quad \begin{aligned} x_{m+1} &= g_1(t), \\ x_{m+2} &= g_2(t), \\ &\cdots\cdots \\ x_n &= g_{n-m}(t) \end{aligned}$$

As long as this should hold, we might be able to express the variables $x_1, x_2, \ldots, x_m$, as functions of $x_{m+1}, x_{m+2}, \ldots, x_n$ in many different ways. For example, it might be possible to express x_1 as a function of x_n say

$$(8.4) \quad x_1 = F(x_n).$$

But could this relation be used to judge the effect upon x_1 of various arbitrary changes in x_n? Obviously not, because the very existence of (8.4) rests upon the assumption that (8.3) holds. The relation (8.4) might be highly unstable for such arbitrary changes, and the eventual persistence observed for (8.4) in the past when (8.3) held good, would not mean anything in this new situation. In the next situation the original system (8.1) or even system (8.2) would still be good, if we knew it. But to find such a basic system of highly autonomous relations in an actual case is not an analytical process, it is a task of making fruitful hypotheses as to how reality actually is.

We shall illustrate these points by two examples.

First we shall consider a scheme which, I think, has some bearing upon the problem of deriving demand curves from time series.

Let x be the rate of per capita consumption of a commodity in a group of people who all have equal money income, R. Let p be the price of the commodity, and let P be an index of cost of living. Assume that the following demand function is actually true:

$$(8.5) \quad x = a\frac{p}{P} + b\frac{R}{P} + c + \varepsilon,$$

where a, b, c, are certain constants, and ε is a random variable with "rather small" variance, and such that the expected values of x are

$$(8.6) \quad E(x|\frac{p}{P}, \frac{R}{P}) = a\frac{p}{P} + b\frac{R}{P} + c.$$

Assume that (8.5) is autonomous in the following sense: For any arbitrary values of p/P and R/P, the corresponding value of x can be estimated by (8.6). Suppose we are interested only in variations that are small relative to certain constant levels of the variables.

Then we may approximate (8.5) by a linear relation in the following way: Let p_o, R_o and P_o be the average values of p, R, and P respectively. Then we have

$$x = a\,\frac{p_o + (p - p_o)}{P_o + (P - P_o)} + b\,\frac{R_o + (R - R_o)}{P_o + (P - P_o)} + c + \varepsilon$$

$$= a\,\frac{p_o + (p - p_o)}{P_o} \cdot \frac{1}{1 + \frac{P - P_o}{P_o}}$$

$$+ b\,\frac{R_o + (R - R_o)}{P_o} \cdot \frac{1}{1 + \frac{P - P_o}{P_o}} + c + \varepsilon$$

(8.5')
$$\cong a\,\frac{p_o + (p - p_o)}{P_o}\left(1 - \frac{P - P_o}{P_o}\right)$$

$$+ b\,\frac{R_o + (R - R_o)}{P_o}\left(1 - \frac{P - P_o}{P_o}\right) + c + \varepsilon$$

$$= \frac{a}{P_o}\,p - \frac{ap_o}{P_o^2}\,P + \frac{ap_o}{P_o} - \frac{a(p - p_o)(P - P_o)}{P_o^2}$$

$$+ \frac{b}{P_o}\,R - \frac{bR_o}{P_o^2}\,P + \frac{bR_o}{P_o} - \frac{b(R - R_o)(P - P_o)}{P_o^2} + c + \varepsilon.$$

If the deviations $(p - p_o)$, $(P - P_o)$, and $(R - R_o)$ are small compared with p_o, P_o, and R_o, we may neglect product terms of these deviations. Then we obtain

(8.7)
$$x = Ap + BR + CP + D + \varepsilon'$$

where
$$A = \frac{a}{P_o}, \quad B = \frac{b}{P_o}, \quad C = -\left(\frac{ap_o}{P_o^2} + \frac{bR_o}{P_o^2}\right), \quad D = \frac{ap_o}{P_o} + \frac{ap_o}{P_o} + \frac{bR_o}{P_o} + c,$$

and where ε' is a new residual term now also containing the errors made by the above approximation. For small variations of the variables, ε' may not be practically distinguishable from ε.

What we shall now show is that, if the data for p, P, and R, to be used for deriving the demand function have, for some reason or another, happened to move as certain regular functions of time, there may in <u>these</u> data exist <u>another</u> relation which has exactly the same <u>form</u> as (8.7), but different coefficients, and which may fit the data still better than

(8.7) would do in general. And if we mistake this other relation for (8.7), we get merely a confluent relationship, and not an approximation to the demand function (8.5).

To see this let us write (8.5) as

$$(8.5'') \qquad x(t) = a\,\frac{p(t)}{P(t)} + b\,\frac{R(t)}{P(t)} + c + \varepsilon(t).$$

Assume now that the time functions $p(t)$, $P(t)$, and $R(t)$ - for some reason - happen to be such that they satisfy the functional relations

$$(8.8) \qquad \frac{p(t)}{P(t)} = k_1 p(t) + k_2 P(t) + k_o$$

$$(8.9) \qquad \frac{R(t)}{P(t)} = m_1 R(t) + m_2 P(t) + m_o,$$

where the k's and the m's are certain constants. A wide class of elementary time functions satisfy such functional equations. And whenever this is the case for the actual observations of p, P, and R, an equation of the form (8.7) could be fitted to the data. But we could not use the equation thus obtained for predicting the effect of an arbitrary price change, or an arbitrary income change, because this equation is not in general an approximation to (8.5) but merely a confluent result of (8.5), (8.8), and (8.9). It, therefore, does not hold, e.g., for price changes which violate (8.8), (8.9), or both.

In general, we have to be very careful in using a particular set of data to modify the form of relationships which we have arrived at on strong theoretical grounds. For example, in the case above we might be led to conclude that (8.7) might be a more correct "form" of the demand function than (8.5), or at least as good, while actually, when (8.8) and (8.9) are fulfilled, we may obtain a relationship of the form (8.7), which is not a demand function at all, and which breaks down as soon as $p(t)$, $P(t)$, and $R(t)$ take on some other time shape.

As an illustration to the question of autonomy of an economic relation with respect to a change in economic policy, let us consider the economic model underlying the famous Wicksellian theory of interest rates and commodity prices. (For the sake of simplicity and shortness we shall, however, make somewhat more restrictive assumptions than Wicksell himself did. Our model does not do full justice to Wicksell's profound ideas.)

Consider a society where there are only three different economic groups: (a) individuals, (b) private firms, and (c) banks. We assume that: (1) All individuals divide their income into two parts, one part consisting of spending+increase in cash-holding, the other part being saved, and all savings go into banks as (time)deposits. There is no other saving in the society. (2) All production in the society takes place in firms. The firms are impersonal organizations, guided in their production policy by profit expectations only. They can make new investments by means of bank loans only. They distribute all their profit to individuals. (3) Prices of goods and services of all kinds vary proportionally through time, and may be represented by a common variable, called the price level. (4) The banks have the power of expanding or contracting credit. We assume that there is only one money rate of interest, which is the same for all banks and the same for loans as for deposits. (This gives a rough description of the model we are going to discuss. It is hardly possible to give an exhaustive description of a model in words. The precise description is given implicitly through the relations imposed in the model).

We are principally interested in the price effect of certain changes in the credit policy of the banks.

Let us introduce the following notations:

(1) $S(t)$=total saving per unit of time,

(2) $I(t)$=total investment per unit of time,

(3) $p(t)$=bank rate of interest at point of time t,

(4) $P(t)$=price level at point of time t,

(5) $R(t)$=total national income per unit of time.

Now we shall introduce a system of fundamental relations describing the mechanism of our model. We consider linear relations, for simplicity.

First, we assume that there exists a market supply function for savings of the following form.

(8.10) $$S(t) = a_0 + a_1 p(t) + a_2 P(t) + a_3 \dot{P}(t) + a_4 R(t).$$

This equation says that the supply of savings(bank deposits) - apart from a constant - depends upon the rate of interest, the total income, the price level, and the expectations regarding the future real value of money saved, as represented by the rate of change in the price level $P(t)$. It might be realistic to assume that a_1 and a_4 are positive, a_2 and a_3 negative.

Next, we assume the following demand function for bank loans:

(8.11) $$I(t) = b_o + b_1 p(t) + b_2 P(t) + b_3 \dot{P}(t),$$

where b_1 is negative and b_3 positive, while the sign of b_2 may be uncertain, a priori. b_3 would be positive because, when the price level is increasing, the firms expect to buy factors of production in a less expensive market than that in which they later sell the finished products, and this profit element is an inducement to invest.

Now, if the banks should lend to firms an amount equal to deposits, neither more nor less, i.e., if

(8.12) $$I(t) = S(t),$$

then it follows from (8.10), (8.11), and (8.12), that to each value of $R(t)$, $P(t)$, and $\dot{P}(t)$, there would correspond a certain market equilibrium rate of interest, $\bar{p}(t)$ called by Wicksell the <u>normal</u> rate. That is, we should have

(8.13) $$\bar{p}(t) = \frac{b_o - a_o}{a_1 - b_1} + \frac{b_2 - a_2}{a_1 - b_1} P(t) + \frac{b_3 - a_3}{a_1 - b_1} \dot{P}(t) - \frac{a_4}{a_1 - b_1} R(t)$$

$$= A_o + A_1 P(t) + A_2 \dot{P}(t) + A_3 R(t),$$

where $\bar{p}(t)$ is a value of $p(t)$ satisfying (8.10), (811), and (8.12), and where the A's are abbreviated notations for the coefficients in the middle term.

If the banks want, actively, to expand or contract currency (that is, if they want to change that amount of money outside the banks), they have to fix a rate of interest $p(t)$ which differs from $\bar{p}(t)$ as defined by (8.13). (Note that $\bar{p}(t)$ is by no means a constant over time.) From (8.10) and (8.11) we get

(8.14) $$I(t) - S(t) = (b_o - a_o) + (b_1 - a_1)p(t) + (b_2 - a_2)P(t) + (b_3 - a_3)\dot{P}(t) - a_4 R(t),$$

which, for $p(t) = \bar{p}(t)$ reduces to

(8.15) $$0 = (b_o - a_o) + (b_1 - a_1)\bar{p}(t) + (b_2 - a_2)P(t) + (b_3 - a_3)\dot{P}(t) - a_4 R(t)$$

Subtracting (8.15) from (8.14) we obtain

(8.16) $$I(t) - S(t) = (b_1 - a_1) [p(t) - \bar{p}(t)]$$

which tells us that the amount of "money inflation," $I(t) - S(t)$, is (negatively) proportional to the difference between the actual bank rate of interest and the normal rate as defined by (8.13).

Assuming the "inflation" stream $I(t) - S(t)$ (taken as a barometer for total spending) to be accompanied by a proportional raise in the price level, we have

(8.17) $\quad \dot{P}(t) = k[I(t) - S(t)]$ $\qquad$ (k a positive constant).

Combining (8.16) and (8.17) we obtain

(8.18) $\quad \dot{P}(t) = k(b_1 - a_1)[p(t) - \bar{p}(t)]$,

which is a simplified expression for Wicksell's fundamental theorem about the price effect of a bank rate of interest that differs from the normal rate.

Accepting this theory (we are not interested in analyzing its actual validity any further in this connection, as we use it merely for illustration), what would be the degree of autonomy of the three equations (8.16), (8.17), and (8.18)?

Let us first consider the equation (8.16). Its validity in our set-up rests upon the two fundamental relations (8.10) and (8.11). In setting up these two equations we did not impose any restrictions upon the time shape of the functions $p(t)$, $P(t)$, and $R(t)$. Therefore, by hypothesis, whatever be the time shape of these functions, the corresponding time shapes of $I(t)$ and $S(t)$ - and, therefore, also the time shape of $I(t) - S(t)$ - follow from (8.10) and (8.11). ((8.16) is merely another way of calculating the difference $I(t) - S(t)$). From (8.13) it follows that to each pair of time functions $P(t)$ (provided its derivative $\dot{P}(t)$ exists) and $R(t)$ there corresponds a time function $\bar{p}(t)$, while to each given time function $\bar{p}(t)$ there corresponds, in general, an infinity of time functions $P(t)$ and $R(t)$. The equation (8.16) is, therefore - by assumption - autonomous in the following sense: For any arbitrarily chosen time functions for $p(t)$ and $\bar{p}(t)$ the credit inflation $I(t) - S(t)$ can be calculated from (8.16).

We should notice that this property of (8.16) - if true - is not a mathematical property of the equation: it cannot be found by looking at the equation. It rests upon a hypothesis as to how the difference $I(t) - S(t)$ in fact would behave for various arbitrary changes in the interest rate and the normal rate. In another model we might obtain an equation of exactly the same form, but without the same property of autonomy. For example, assume that - as a consequence of some model, whatever be the particular economic reasoning underlying it - all the time functions above were bound to follow certain linear trends. In particular, suppose that we had $I(t) - S(t) = mt$, $p(t) - \bar{p}(t) = nt$. We should then have

(8.19) $\quad I(t) - S(t) = \frac{m}{n}[p(t) - \bar{p}(t)]$

which is of the form (8.16). But from (8.19) we could not calculate the effect upon $I(t) - S(t)$ of say, various types of interest policy, because any changes in $p(t)$ that would violate the condition $p(t) - \bar{p}(t) = nt$ would break up the very foundation upon which (8.19) rests. The equation (8.19) might still hold after such a break, but that would have to follow from another model.

The equation (8.17) represents, per se, also an autonomous relation with respect to certain changes in structure. It is an independent hypothesis about the price level saying that, whatever be the credit inflation $I(t) - S(t)$, we may calculate the corresponding rate of change in the price level. Here too, we cannot know how far this property of autonomy could in fact be true. It is an assumption, and it is a task of economic theory and research to justify it.

Let it be established that (8.16) and (8.17) are, in fact, highly autonomous relations. What is the situation with respect to the equation (8.18)? Obviously (8.18) would have a smaller degree of autonomy than either (8.16) or (8.17) separately, because the class of time functions satisfying (8.18) is - by definition - only the class of functions that satisfy (8.16) and (8.17) jointly.

So far we have not assumed any definite relations describing the credit policy of the banks. We have merely described the behavior of individuals and firms in response to a given bank rate of interest. Starting from certain assumptions as to the willingness to save and to invest, and assuming that an inflow of extra credit into the market causes a proportional change in the price level, we have obtained two structural relations (8.16) and (8.17). The variable $p(t)$ was considered as a free parameter. It might be, however, that the banks, over a certain period of time at least, choose to follow a certain pattern in their interest policy, or that they have to do so in order to secure their own liquidity. Over this period of time it might then be that we could add a new relation to the ones above, namely a relation describing - temporarily - the banking policy. Assume for instance, that the banks, over a certain period of time, act as follows: Whenever they realize that $I(t) - S(t)$ has become positive they start raising the interest rate, in order to protect their liquidity, and, conversely, they lower the rate of interest when they realize a negative $I(t) - S(t)$. Such a policy might be described by the relation

(8.20) $$\dot{p}(t) = c[I(t) - S(t)],$$

where c is a positive constant. Because of (8.16) we have

(8.21) $\dot{p}(t) = c(b_1 - a_1)[p(t) - \bar{p}(t)]$

And combining (8.18) and (8.21) we have

(8.22) $\dot{P}(t) = \frac{k}{c}\dot{p}(t)$

which apparently says that the price level moves in the <u>same</u> direction as the interest rate. But could we use this relation to calculate the "would-be" effect upon the price level of some arbitrary interest policy? Obviously <u>not</u>, because (8.22) holds only when $R(t)$, $I(t)$, $S(t)$, $P(t)$, $p(t)$, and $\bar{p}(t)$ are such time functions as satisfy, simultaneously, (8.13), (8.16), (8.17), and (8.20). Therefore, (8.22) is of no use for judging the effect of a change in interest policy. To obtain an equation for this purpose we might combine (8.13) and (8.18), which give a relation of the form

(8.23) $\dot{P}(t) + BP(t) = H_1 p(t) + H_2 R(t) + H_0$,

where B, H_1, H_2, and H_0 are constants depending upon those in (8.13) and (8.18). Here there are - by hypothesis - no restrictions upon the time shape of the functions $p(t)$ and $R(t)$. We may choose such functions arbitrarily and solve the equation (8.23) to obtain $P(t)$ as an explicit function of $p(t)$ and $R(t)$.

But how could we know that (8.23) is the equation to use, and not (8.22)? There is no formal method by which to establish such a conclusion. In fact, by starting from another model with different assumptions, we might reach the opposite conclusion. To reach a decision we have to <u>know</u> or to imagine - on the basis of general experience - which of the two relations (8.22) or (8.23) would <u>in fact</u> be the most stable one if either of them were used as an autonomous relation.

* * *

To summarize this discussion on the problem of autonomous relations: In scientific research - in the field of economics as well as in other fields - our search for "explanations" consists of digging down to more fundamental relations than those that appear before us when we merely "stand and look". Each of these fundamental relations we conceive of as invariant with respect to a much wider class of variations than those particular ones that are displayed before us in the natural course of events. Now, if the real phenomena we observe day by day are really ruled by the simultaneous action of a whole system of fundamental laws, we see only very little of the whole class of hypothetical variations for which each of the fundamental relations might be assumed to hold. (This fact also raises very serious problems of estimating fundamental relations from current observations. This whole problem we shall discuss in Chapter V.) For the variations we observe, it is possible to establish an infinity of relationships, simply by combining two or more of the fundamental relations in various ways. In particular, it might be possible to write one economic variable as a function of a set of other variables, in a great variety of ways. To state, therefore, that an economic variable is "some function" of a certain set of other variables, does not mean much, unless we specify in what "milieu" the relation is supposed to hold. This, of course, is just another aspect of the general rule we laid down at the beginning of this chapter: The rule that every theory should be accompanied by a design of experiments.

4. Parts of Professor Frisch's article "Repercussion Studies at Oslo." American Economic Review. Volume XXXVIII, No. 3, June, 1948.

A third feature of the work, and the most important, is the analysis of the structural relations that connect the various variables in the national budgeting system. The national budgeting system itself is only a framework to keep track of the variables involved. The essence of the problem, that part of it in which the questions that are really significant from the repercussion viewpoint emerge, is this system of structural relations. A few examples will suffice to indicate what I have in mind when I speak of structural relations. Take, for instance, the relations that show how the consumption of a certain commodity depends on the price and on the income in the consumer group and on some other factors. Or, take the relations that show how the expectations of business enterprises determine the investment activity. These ideas were studied intensively in the fundamental work of Knut Wicksell some forty years ago and have in our day been brought into popularity through the writings of Keynes. As a further example, I may mention, on the one hand, the connection that exists between multiplier and acceleration effects and on the other, the way in which the time curve of the starting of capital production determines the time curve of the carry-on-activity in the capital producing industries. This type of problem was studied in an essay on impulse problems and propagation problems in economic dynamics which I published some years before the war in the volume in honour of Cassel. Further, we have the important group of problems centering around the supply structure of labour. These examples will be enough to indicate what I have in mind when I speak of structural relations.

From the viewpoint of repercussion studies, it is essential to look somewhat closer into the nature of these structural relations. For one thing, it must be pointed out that a structural relation is conceived of as one in which the variables are capable of a partial variation. That is to say, the relation is conceived of as expressing what would happen if all the variables involved were kept constant except two, and these two varied together. Further, we must consider how autonomous the relation is. This point is often neglected. From the viewpoint of repercussion studies, however, it becomes of crucial importance. The idea of autonomy can be explained as follows. Take any equation and ask the question: is the technical and institutional setting which surrounds it and the behaviour of the individuals involved such that this particular equation will hold good even though other equations involving the same variables are destroyed through technical, institutional or behaviouristic changes or through the fixation of some specific variables in the system, for instance through a specific economic measure. This, it seems, is the only way in which it is possible to define a "causal" relation as distinguished from an incidental covariation between economic magnitudes.

Obviously, this feature of an equation is an exceedingly important one when a question of repercussion comes up. The nature of an equation in this respect should if possible be discussed in quantitative terms. Some sort of coefficient should be attributed to it expressing its degree of autonomy. Similar coefficients may be applied to parts of the equation. Without going into detail I may say this much, that the autonomy coefficients will have properties to a large extent similar to those of probabilities. They will in a sense express the probability that the relation in question exists under certain assumptions regarding the existence or non-existence of other equations, or parts of equations. The smaller the range of assumption needed in order to assure a certain level of probability, the larger the autonomy of the equation, or part of equation, in question. Coefficients of this sort will be equally important I believe in a planned economy and in an economy that is based essentially on free enterprise.

The question then is, how can we proceed in order to arrive at a workable system of structural equations? In the first place, a considerable amount of speculative thinking must be done. We must use both the theoretical approach and as much common sense as we can.

. . . .

These equations will then be scrutinized very thoroughly, particularly with a view to throwing out any equation that does not have a high degree of autonomy but is only an expression for an incidental covariation.

The further we can go in the direction of eliminating these incidental equations from the system and getting down to equations of as high a degree of autonomy as possible, the more sure we will be of having a system which we can rely upon when we are going to answer questions regarding repercussions, *i.e.*, questions which will in certain cases violate some of the equations in our system, or change some of the structural coefficients in these equations.

Having through this *a priori* approach arrived at a preliminary conclusion regarding a desirable system of relations to be considered, the question arises how to get information about the *numerical* character of the equations, *i.e.*, information about the magnitude of the coefficients entering into them. This raises a difficult question which is usually not considered as carefully as it merits.

Let me give you a simplified example of the difficulty I have in mind. Suppose you are considering a system of three variables. You may measure them along the axes in a three-dimensional space. A set of observations of these variables would be represented by a scatter of points in this three-dimensional space. Suppose you are trying to analyze a demand relation, one of the three variables standing for the quantity of a certain commodity, the second standing for its price, and the third for the income of the consumers. This is tantamount to assuming that, apart from random disturbances, your scatter will be distributed along a certain surface in the three-dimensional space. To try to determine the numerical character of this demand relation from the statistical observations would mean that you let the shape of the scatter of points determine the coefficients of the equation. But now suppose that the statistical data at hand satisfy also some other relation, for instance, a supply relation. That would mean that the scatter of points not only lies in the first surface but also in a second surface, the supply surface. In other words, the scatter of points would not be spread out over the demand surface but would, apart from random disturbances, be concentrated along a specific curve in the demand surface. Obviously in this case, it would be impossible from the nature of the scatter to get complete information about the nature of the demand surface. In other words, we would not have any information about demand in the sense of a relation that expresses what would take place if the variables involved underwent a partial variation. Thus, if by any chance the data at hand satisfy not only that specific relation which you happen to think of at the moment and on which you try to get numerical information, but also satisfy some other relation, then it is in point of principle impossible to get the kind of numerical information you are after. You will see immediately that in a great number of cases, indeed in most cases, the data which usually present themselves to us in economic statistics are of just this kind. It is very seldom indeed that we have a clear case where the statistical data can actually determine numerically an autonomous structural equation. In most cases we only get a covariational equation with a low degree of autonomy.

This situation is extremely important from the repercussion viewpoint. We must look for some other means of getting information about the numerical character of our structural equations. The only possible way seems to be to utilize to a much larger extent than we have done so far the interview method, i.e., we must ask persons or groups what they would do under such and such circumstances. In doing so we must, of course, watch our step very carefully to avoid bias in the answers. A number of pitfalls exist in this field. But, after all, I think that the possibility of getting reasonably reliable answers is better than would appear at first sight. There are two things we may do in order to assure the answers to be as reliable as possible. In the first place, we may use questions which are worked out in such a way that the information we seek does not come directly from the answers themselves but rather from the solution of a system of equations connecting these answers, i.e., we try to conceal as much as possible from the interviewed persons or groups the true object of the interviewing. In the second place, we will have a check on the answers by noticing whether or not the relations we derive from them, check with the covariational equations which we derive from the usual kind of statistics. Proceeding along such lines it should be possible to work out a system of corrections to be applied to the answers so as to adjust for most of the bias. Of course it is impossible to state any general rule concerning this. Each particular field of enquiry must be judged on its own merits and carefully studied.

5. "Identification Problems in Economic Model Construction" by T.C. Koopmans. Cowles Commission Discussion Papers, Statistics: 316. July 19, 1948.

1. Statistical inference and the construction of models. In recent years an increasing number of authors has resorted to the construction of economic models as the principal tool of the analysis of economic fluctuations and related problems of policy. In these models, macro-economic variables are thought of as determined by a complete system of equations. The meaning of the term "complete" is discussed more fully below. At present it may suffice to describe a complete system as one in which there are as many equations as endogenous variables, that is, variables whose formation is to be "explained" by the equations. The equations are usually of three kinds: equations of economic behavior, technological laws of transformation, and definitions. We shall use the term structural equations to comprise all three types of equations.

Systems of structural equations may be composed entirely on the basis of economic "theory". By this term we shall understand the combination of (a) principles of economic behavior derived from general observation - partly introspective, partly through interview or experience - of the motives of economic decisions, (b) technological knowledge, and (c) carefully constructed definitions of variables. Alternatively, a structural equation system may be determined on the dual basis of such "theory" combined with systematically collected statistical data for the relevant variables for a given period and country or other unit. In this article we shall discuss certain problems that arise out of model construction in the second case.

Where statistical data are used as one of the foundation stones on which the equation system is erected, the modern theory and methods of statistical inference are an indispensable instrument. However, without "theory" as another foundation stone, it is impossible to make such statistical inference apply directly to the equations of economic behavior which are most relevant to analysis and to policy discussion. Statistical inference unsupported by economic theory applies to whatever statistical regularities and stable relationships can be discerned in the data.* Such purely empirical relationships when discernable are likely to be due to the presence and persistence of the underlying structural relationships, and (if so) could be deduced from a knowledge of the latter. However, the direction of this deduction cannot be reversed - from the empirical to the structural relationships - except possibly with the help of a theory which specifies the form of the structural relationships, the variables which enter into each, and any further details supported by prior observation or deduction therefrom. The more detailed these specifications are made in the model, the greater scope is thereby given to statistical inference from the data to the structural equations. We propose to study the limits to which statistical inference, from the data to the structural equations (other than definitions), is subject, and the manner in which these limits depend on the support received from economic theory.

This problem has attracted recurrent discussion in econometric literature, with varying terminology and degree of abstraction. Reference is made to Pigou (15) , Henry Schultz, (16, especially Chapter II, Section IIIe), Frisch (4,5), Marschak (14, especially Sections IV and V), Haavelmo (6, especially Chapter V). An attempt to systematize the terminology and to formalize the treatment of the problem has been made over the past few years by various authors connected in one way or another with the Cowles Commission for Research in Economics. Since the purpose of this article is expository, I shall draw freely on the work by Koopmans and Rubin 13 , Wald 17 , Hurwicz 7 , Koopmans, Rasch and Reiersøl (12), without specific acknowledgement in each case. We shall proceed by discussing a sequence of examples, all drawn from econometrics, rather than by a formal logical presentation, which can be found in references (7) and (12).

2. Concepts and examples. The first example, already frequently discussed, is that of a competitive market for a single commodity, of which the price p and the quantity q are determined through the intersection of two rectilinear schedules, of demand and supply respectively, with instantaneous response of quantity to price in both cases. For definiteness' sake, we shall think of observations as applying to successive periods in time. We shall further assume that the slope coefficients α and γ of the demand and supply schedules respectively are constant through time, but that the levels of the two schedules are subject to not directly observable shifts from an equilibrium level. The structural equations can then be written as:

* See T.C. Koopmans (11).

$$(1) \quad \begin{cases} (1d)\ q + \alpha p + \varepsilon = u & \text{(demand)} \\ (1s)\ q + \gamma p + \eta = v & \text{(supply)} \end{cases}$$

Concerning the shift variables u and v we shall assume that they are random drawings from a stable joint probability distribution with mean values equal to zero:

$$(2) \qquad \phi(u,v), \ \mathcal{E}u = 0, \ \mathcal{E}v = 0.$$

We shall introduce a few terms which we shall use with corresponding meaning in all examples. The not directly observable shift variables u,v are called latent variables, as distinct from the observed variables, p,q. We shall further distinguish structure and model. By a structure we mean the combination of a specific set of structural equations (1) (such as is obtained by giving specific numerical values to $\alpha, \gamma, \varepsilon, \eta$), and a specific distribution function (2) of the latent variables (for instance a normal distribution with specific, numerically given, variances and covariance). By a model we mean only a specification of the form of the structural equations (for instance their linearity and a designation of the variables occurring in each equation), and of a class of functions to which the distribution function of the latent variables belongs (for instance, the class of all normal bivariate distributions with zero means). More abstractly, a model can be defined as a set of structures. For a useful analysis, the model will be chosen so as to incorporate relevant *a priori* knowledge or hypotheses as to the economic behavior to be described. For instance, the model here discussed can often be narrowed down by the usual specification of a downward sloping demand curve and an upward sloping supply curve:

$$(3) \qquad \alpha > 0, \ \gamma < 0.$$

Let us assume for the sake of argument that the observations are produced by a structure, to be called the "true" structure, which is contained in (permitted by) the model. In order to exclude all questions of sampling variability (which are a matter for later separate inquiry), let us further make the unrealistic assumption that the number of observations produced by this structure can be increased indefinitely. What inferences can be drawn from these observations toward the "true" structure?

A simple reflection shows that in our present example neither the "true" demand schedule nor the "true" supply schedule can be determined from any number of observations. To put the matter geometrically, let each of the two identical scatter diagrams in figures

1A and 1B represent the jointly observed values of p and q. A structure compatible with these observations can be obtained as follows: Select arbitrarily "presumptive" slope coefficients α and γ of the demand and supply schedules. Through each point $S(p,q)$ of the scatter diagrams draw two straight lines with slopes given by these coefficients. The presumptive demand and supply schedules will intersect the quantity axis at distances $-\varepsilon + u$ and $-\eta + v$ from the origin, <u>provided</u> the presumptive slope coefficients α and γ are the "true" ones. We shall assume this to be true in figure 1A. In that case the values of ε and η can be found from the consideration that the averages of u and v in a sufficiently large sample of observations are practically equal to zero.

However, nothing in the situation considered permits us to distinguish the "true" slopes α, γ (as shown in Figure 1A) from any other presumptive slopes (as illustrated in Figure 1B).

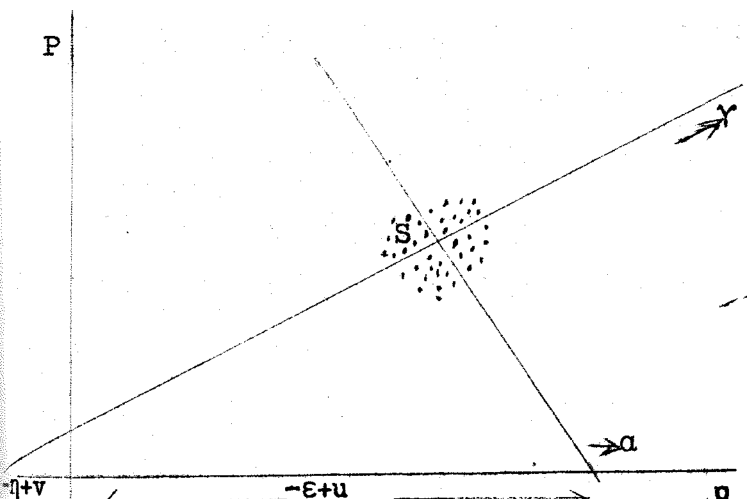

Figure IA

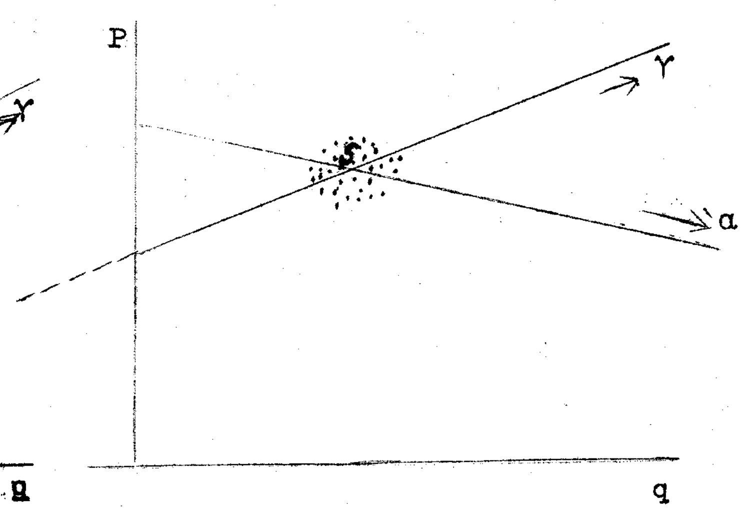

Figure IB

Any arbitrary set of slope coefficients represents another, statistically just as acceptable, hypothesis concerning the formation of the observed variables.

Let us formulate the same remark algebraically in preparation for further examples in more dimensions. Let the numerical values of the "true" parameters α, γ, ε, η. in (1) be known to an individual who, taking delight in fraud, multiplies the demand equation (1d) by 2/3, the supply equation (1s) by 1/3, and adds the result to form an equation

$$(2d) \qquad q + \frac{2\alpha + \gamma}{3} p + \frac{2\varepsilon + \eta}{3} = u'$$

which he proclaims to be the demand equation. This equation is actually different from the "true" demand equation (1d) because (3) implies $\alpha \neq \gamma$. Similarly he multiplies the same equations by 2/5 and 3/5 respectively, say, to produce an equation

$$\text{(2s)} \qquad q + \frac{2\alpha + 3\gamma}{5}\, p + \frac{2\varepsilon + 3\eta}{5} = v'$$

different from the "true" supply equation (1s), but which he presents as if it were the supply equation. If our prankster takes care to select his multipliers in such a manner as not to violate the sign rules (3) imposed by the model, the deceit cannot be discovered by statistical analysis of any number of observations.* For the equations (2), being derived from (1), are satisfied by all data that satisfy the "true" equations(1). Moreover, being of the same form as the equations (1), the equations (2) are equally acceptable a priori.

The second example differs from the first only in that the model specifies a supply equation containing in addition an exogenous variable. To be definite, we shall think of the supply of an agricultural product as affected by the rainfall r during a critical period of crop growth** or crop gathering. This variable is called exogenous to our model to express the plausible hypothesis that rainfall v, while affecting the market of the commodity concerned, is not itself affected thereby. Put in mathematical terms, this hypothesis specifies that the disturbances u and v in

$$(3) \quad \begin{cases} \text{(3d)} \quad q + \alpha p \qquad\quad + \varepsilon = u & \text{(demand)} \\ \text{(3s)} \quad q + \gamma p + \delta v + \eta = v & \text{(supply)} \end{cases}$$

are statistically independent*** of the values assumed by v.

---

* The deceit could be discovered if the model were to specify a property (e.g., independence) of the disturbances u and v which is not shared by $u' = (2u + v)/3$ and $v' = (2u + 3v)/5$. We have not made such a specification.

** With respect to this example, the assumption of a linear relationship can be maintained only if we think of a certain limited range of variation in rainfall. Another difficulty with the example is that for most agricultural products, the effect of price on supply is delayed instead of instantaneous, as here assumed. A practically instantaneous effect can, however, be expected in the gathering of wild fruits of nature.

*** It is immaterial for this definition whether the exogenous variable is regarded as a given function of time - a concept perhaps applicable to a variable set by government policy - or as itself a random variable determined by some other structure involving probability distributions - a concept applicable particularly to weather variables. It should further be noted that we postulate independence between v and (u,v), not between v and (p,q), although we wish to express that v "is not affected by" p and q. The meaning to be given to the latter phrase is that in other equations explaining the formation of v the variables (p,q) do not enter. Precisely this is implied in the statistical independence of v and (u,v), because (p,q) is by virtue of (3) statistically dependent on (u,v), and any role of (p,q) in the determination of v would therefore create statistical dependence between v and (u,v). On the other hand, the postulated statistical independence between v and (u,v) is entirely compatible with the obvious influence, by virtue of (3), of v on (p,q).

It will be seen at a glance that the supply equation can still not be determined from a sample of any size. If, starting from "true" structural equations (3) we multiply by -1/2 and 3/2, say, and add the results to obtain a pretended supply equation,

$$(4s) \qquad q + \frac{3\gamma - \alpha p}{2} p + \frac{3\delta}{2} v + \frac{3\eta - \varepsilon}{2} = v'$$

of the same prescribed form as (3s), any data will satisfy this equation (4s) as well as they satisfy the two equations (3).

A similar reasoning can <u>not</u> be applied to the demand equation in the present model. Any attempt to construct another pretended demand equation by a linear combination involving the supply equation (3s) would introduce into that pretended demand equation the variable v which by the hypotheses underlying the model does not belong in it.

It might be thought that, if v has the properties of a random variable, its presence in the pretended demand equation might be concealed because its "contribution" cannot be distinguished from the random disturbance in that equation. To be specific, if 4/3 and and -1/3 are arbitrarily selected multipliers, the disturbance in the pretended demand equation might be thought to take the form

$$(5) \qquad u' = \frac{4u - v}{3} \frac{-1}{3} v.$$

This, however, would violate the specification that v is exogenous and that therefore v and u' are to be statistically independent as well as v and (u,v). The relevance of the exogenous character of v to our present discussion is clearly illustrated by this remark.

Our analysis of the second example suggests (and below we shall cite a theorem establishing proof) that a sufficiently large sample does indeed contain information with regard to the parameters α, ε of the demand equation - it being understood that such information is conditional upon the validity of the model. It can already now be seen that there must be the following exception to this statement. If in fact (although the model does not require it) rainfall has no influence on supply, that is, if in the "true" structure $\delta = 0$, then any number of observations must necessarily be compatible with the model (1), and hence does not convey information with regard to either the demand equation or the supply equation.

As a <u>third example</u> we consider a model obtained from the preceding one by the inclusion in the demand equation of consumers' income i as an additional exogenous variable. We assume the exogenous character of consumers' income merely for reasons of exposition, and in full awareness of the fact that actually price and quantity on any market do affect income directly to some extent, while furthermore the disturbances u and v affecting the market under consideration may well be correlated with similar disturbances in several other markets which together have a considerably larger effect on consumers' income.

The structural equations are now

$$(6) \quad \begin{cases} (6d) & q + \alpha p + \beta i + \varepsilon = u \quad \text{(demand)} \\ (6s) & q + \gamma p + \delta v + \eta = v \quad \text{(supply)} \end{cases}$$

Since each of the two equations now excludes a variable specified for the other equation, neither of them can be replaced by a different linear combination of the two without altering its form. This suggests, and proof is cited below, that from a sufficiently large sample of observations, the demand equation can be determined provided rainfall actually affects supply $(\alpha \neq 0)$, and the supply equation can be determined provided consumers' income actually affects demand $(\beta \neq 0)$.

The <u>fourth example</u> is designed to show that situations may occur in which some but not all parameters of a structural equation can be determined from sufficiently many observations. Let the demand equation contain both this year's income i_0 and last year's income i_{-1} but let the supply equation not contain any variable absent from the demand equation:

$$(7) \quad \begin{cases} (7d) & q + \alpha p + \beta_0 i_0 + \beta_{-1} i_{-1} + \varepsilon = u \\ (7s) & q + \gamma p + \eta = v \end{cases}$$

Now obviously we cannot determine either α or ε, because linear combinations of the equations (7) can be constructed which have the same form as (7d) but other* values α' and ε' for the coefficients α and ε. However, as long as (7d) enters with some non-vanishing weight into such a linear combination, the ratio

$$(8) \qquad \beta_{-1} | \beta_0$$

is not affected by the substitution of that linear combination for the "true" demand equation. Thus, if the present model is correct, the observations contain information with respect to the relative importance of present and past income to demand, whereas they are silent on the price elasticity of demand.

* As regards ε' this is true whenever $\varepsilon \neq \eta$. As regards α' it is safe-guarded by (3).

The fifth example shows that an assumption regarding the joint distribution of the disturbances u and v, where justified, may open the door to a determination of a structural equation which is otherwise indeterminate. Returning to the equation system (3) of our second example, we shall now make the model specify in addition that the disturbances u in demand and v in supply are statistically independent. Remembering our previous statement that the demand equation can already be determined without the help of such an assumption, it is clear that in attempting to construct a "pretended" supply equation, no linear combination of the "true" demand and supply equations (3), other than the "true" supply equation (3s) itself, can be found which preserves the required independence of disturbances in the two equations. Writing λ and $1-\lambda$ for the multipliers used in forming such a linear combination, the disturbance in the pretended supply equation would be

$$(9) \qquad v' = \lambda u + (1 - \lambda)v.$$

Since u and v are by assumption independent, the disturbance v' of the pretended supply equation is independent of the disturbance u in the demand equation already found determinable if and only if $\lambda = 0$, i.e. if the pretended supply equation coincides with the "true" one.

3. The identification of structural parameters. In our discussion we have used the phrase "a parameter that can be determined from a sufficient number of observations". We shall now define this concept more sharply, and give it the name identifiability of a parameter. Instead of reasoning, as before, from "a sufficiently large number of observations" we shall base our discussion on a hypothetical knowledge of the probability distribution of the observations, as defined more fully below. It is clear that exact knowledge of this probability distribution cannot be derived from any finite number of observations. Such knowledge is the limit approachable but not attainable by extended observation. By nevertheless hypothesizing the full availability of such knowledge, we obtain a clear separation between problems of statistical inference arising from the variability of finite samples, and problems of identification in which we explore the limits to which inference even from an infinite number of observations is subject.

A structure has been defined as the combination of a distribution of latent variables and a complete set of structural equations. By a complete set of equations we mean a set of as many equations as there are endogenous variables. Each endogenous variable may occur with or without time lags, and should occur without lag in at least one equation.

Finally, the set should be such as to permit unique determination of the non-lagged values of the endogenous variables from those of the lagged endogenous, the exogenous, and the latent variables. By endogenous variables we mean observed variables which are not exogenous, i.e., variables which are not known or assumed to be statistically independent of the latent variables, and whose occurrence in one or more equations of the set is necessary for its completeness.

It follows from these definitions that, for any specific set of values of the exogenous variables, the distribution of the latent variables (one of the two components of a given structure) entails or generates, through the structural equations (the other component of the given structure), a probability distribution of the endogenous variables. The latter distribution is, of course, conditional upon the specified values of the exogenous variables for each time point of observation. This conditional distribution, regarded again as a function of all specified values of exogenous variables, shall be the hypothetical datum for our discussion of identification problems.

We shall call two structures S and S' (observationally) equivalent (or indistinguishable) if the two conditional distributions of endogenous variables generated by S and S' are identical for all possible values of the exogenous variables. We shall call a structure S permitted by the model (uniquely) identifiable within that model if there is no other equivalent structure S' contained in the model. Although the proof has not yet been completely indicated, it may be stated in illustration that in our third example almost all structures permitted by the model are identifiable. The only exceptions are those with either $\beta = 0$ or $S = 0$ (or both). In the first and second examples, however, no structure is identifiable, although in the second example, we have stated that the demand equation by itself is determinate. To cover such cases we shall say that a certain parameter θ of a structure S is uniquely identifiable within a model, if that parameter has the same value for all structures S' equivalent to S, contained in the model. Finally, a structural equation is said to be identifiable if all its parameters are.

This completes the formal definitions with which we shall operate. They can be summarized in the statement that anything is called identifiable, the knowledge of which is implied in the knowledge of the distribution of the endogenous variables, accepting the model as valid. We now proceed to a discussion of the application of this consept to linear models of the kind illustrated by our examples.

4. Identifiability criteria in linear models. In our discussion of these examples, it has been possible to conclude to non-identifiability of certain structural equations or parameters, whenever we were able to construct different linear combinations of some or all equations, which likewise meet the specifications of the model. In the opposite case, where we could show that no such different linear combinations exist, we could not yet conclude definitely that the equation involved is identifiable. Could perhaps other operations than linear combination produce equations of the same form?

We shall now cite a theorem which establishes that no other operations can achieve this. The theorem relates to models specifying a complete set of structural equations as defined above, and in which endogenous and exogenous variables enter linearly. Any time lags with which these variables may occur are supposed to be integral multiples of the time unit to which each observation applies. Furthermore the exogenous variables (considered as different variables whenever they occur with a different time lag) are assumed to be linearly independent. Finally, although simultaneous disturbances in different structural equations are permitted to be correlated, it is assumed that any disturbances operating i different time units (whether in the same or in different structural equations) are statistically independent.

Suppose the model does not specify anything beyond what has been stated. That is, no restrictions are specified yet, that exclude some of the variables from specific equations. Obviously, with respect to such a broad model, not a single structural equation is identifiable. However, a theorem has been proved (13) to the effect that, given a structure S within that model, any structure S' in the model, equivalent to S, can be derived from S by replacing each equation by some linear combination of some or all equations of S.

It will be clear that this theorem remains true if the model is narrowed down by excluding certain variables from certain equations, or by other restrictions on the parameters. Thus, whenever in our examples we have concluded that different linear combinations of the same form prescribed for a structural equation did not exist, we have therewith established the identifiability of that equation. More in general, the analysis of the identifiability of a structural equation in a linear model consists in a study of the possibility to produce a different equation of the same prescribed form by linear combination of all equations. If this is shown to be impossible, the equation in question is thereby proved to be identifiable. To find criteria for the identifiability of a structural equation in a linear model is therefore a straightforward mathematical problem, to which the solution has been given elsewhere(13). Here we shall state without proof what the criteria are.

A necessary condition for the identifiability of a structural equation within a given linear model is that the number* of variables excluded from that equation (more generally: the number of linear restrictions on the parameters of that equation) be at least equal to the number (G, say) of structural equations less one. This is known as the order condition of identifiability. A necessary and sufficient condition for the identifiability of a structural equation within a linear model, restricted only by the exclusion of certain variables from certain equations, is that we can form at least one non-vanishing determinant of order G-1 out of those coefficients, properly arranged, with which the variables, excluded from that structural equation, appear in the G-1 other structural equations. This is known as the rank condition of identifiability.

The application of these criteria to the foregoing examples is straightforward. In all cases considered, the number of structural equations is $G = 2$. Therefore, any of the equations involved can be identifiable only if at least $G - 1 = 1$ variable is excluded from it by the model. If this is so, the equation is identifiable provided at least one of the variables so excluded occurs in the other equation with non-vanishing coefficient (a determinant of order 1 equals the value of its one and only element).

5. Identification through disaggregation. As a further exercise in the application of these criteria, we shall consider a question which has already been the subject of a discussion between Ezekiel (2,3) and Klein (8,9). The question is whether identifiability of the investment equation can be attained by the subdivision of the investment variable into separate categories of investment. In the discussion referred to, which took place before the concepts and terminology employed in this article were developed, questions of identifiability were discussed alongside with questions regarding the merit of particular economic assumptions incorporated in the model, and with questions of the statistical method of estimating parameters that have been recognized as identifiable. In the present context, we shall avoid the latter two groups of problems and concentrate on the formal analysis of identifiability, accepting a certain model as valid for purposes of discussion.

As a starting point we shall consider a simple model expressing the crudest elements of Keynesian theory. The variables are, in money amounts,

---

*Again counting lagged variables as separate variables.

$$(10)\quad \begin{cases} S & \text{savings} \\ I & \text{investment} \\ Y & \text{income} \\ Y_{-1} & \text{income lagged one year.} \end{cases}$$

The structural equations are:

$$(11)\quad \begin{cases} (11id) & S - I = 0 \\ (11S) & S - \alpha_1 Y - \alpha_2 Y_{-1} - \alpha_0 = u \\ (11I) & I - \beta_1 Y - \beta_2 Y_{-1} - \beta_0 = v \end{cases}$$

Of these, the first is a definition expressing the well-known savings-investment identity. The second is a behavior equation of consumers, indicating that the money amount of their savings (income not spent for consumption) is determined by present and past income, subject to a random disturbance u. The third is a behavior equation of entrepreneurs, indicating that the money amount of investment is determined by present and past income, subject to a random disturbance v.

Since the identity (11d) is fully given a priori, no question of identifiability arises with respect to the first equation. In both the second and third equations, only one variable is excluded which appears in another equation of the model, and no other restrictions on the coefficients are stated.* Hence both of these equations already fail to meet the necessary order criterion of identifiability. This could be expected because the two equations connect the same savings-investment variable with the same two income variables, and can therefore not be distinguished statistically.

Ezekiel attempts to obtain identifiability of the structure by a refinement of the model through subdivision of aggregate investment I in the following four components:

(12a)
- I_1 investment in plant and equipment
- I_2 investment in housing
- I_3 temporary investment: changes in consumers' credit and in business inventories.
- I_4 quasi-investment: net contributions from foreign trade and the government budget.

* The normalization requirement that the variables S and I shall have coefficients +1 in (11S) and (11I) respectively does not restrict the relationships involved but merely serves to give a common level to coefficients which otherwise would be subject to arbitrary proportional variation.

For each of these types of investment decisions, a separate explanatory equation is either introduced explicitly or implied in the verbal comments. In attempting to formulate these explanations in terms of a complete set of behavior equations we shall introduce two more variables:

$$(12b)\quad \begin{cases} H & \text{semi-independent cycle in housing investment} \\ E & \text{autonomous component of quasi-investment.} \end{cases}$$

In addition, linear and quadratic functions of time are introduced as trend terms in some equations by Ezekiel. For purposes of the present discussion, we may as well disregard such trend terms, because they would help toward identification only if they could be excluded *a priori* from some of the equations while being included in others - a position advocated neither by Ezekiel nor by the present author.

With these qualifications "Ezekiel's model" can be interpreted as follows:

$$\begin{cases} (13id) & S - I_1 - I_2 - I_3 - I_4 = 0 \\ (13S) & S - \alpha_1 Y - \alpha_2 Y_{-1} - \alpha_0 = u \\ (13I_1) & I_1 - \beta_1 Y - \beta_2 Y_{-1} - \beta_0 = v_1 \\ (13I_2) & I_2 - \gamma_1 Y - \gamma_2 Y_{-1} - H - \gamma_0 = v_2 \\ (13I_3) & I_3 - \delta_1 Y - \delta_2 Y_{-1} - \delta_0 = v_3 \\ (13I_4) & I_4 - \varepsilon_1 Y - \varepsilon_2 Y_{-1} - E - \varepsilon_0 = v_4 \end{cases} \qquad (13)$$

(13id) is the savings-investment identity. (13S) repeats (11S), and ($13I_1$) is modeled after (11I). More specific explanations are introduced for the three remaining types of investment decisions.

Housing investment decisions I_2 are explained partly on the basis of income* Y, partly on the basis of a "semi-independent housing cycle" H. In Ezekiel's treatment H is not an independently observed variable, but a smooth long cycle fitted to I. We share Klein's objection (8, p. 255) to this procedure, but do not think that his proposal to substitute a linear function of time for H does justice to Ezekiel's argument. The latter definitely thinks of H as produced largely by a long-cycle mechanism peculiar to the housing market, and quotes in support of this view a study by Derksen (1) in which this mechanism is analyzed. Derksen constructs an equation explaining residential con-

* We have added a term with Y_{-1} because the exclusion of such a term could hardly be made the basis for a claim of identifiability.

struction in terms of the rent level, the rate of change of income, the level of building cost in the recent past, and growth in the number of families; he further explains the rent level in terms of income, the number of families and the stock of dwelling units (all of these subject to substantial time lags). The stock of dwelling units, in its turn, represents an accumulation of past construction diminished by depreciation or demolition. Again accepting without inquiry the economic assumptions involved in these explanations, the point to be made is that H in ($13I_2$) can be thought to represent specific observable exogenous and past endogenous variables.

Temporary investment I_3 is related by Ezekiel to the rate of change in income. Quasi-investment I_4 is related by him partly to income* (especially via government revenue, imports), partly to exogenous factors underlying exports and goverment expenditure where used as an instrument of policy. The variable E in ($13I_4$) is therefore similar to H in that it can be thought to represent observable exogenous or past endogenous variables. We conclude that the presence of the variables H and E so interpreted does not upset the completeness of the set of equations (13) in the sense defined above.

Let us now apply our criteria of identifiability to the behavior equations in (13). In each of these, the number of excluded variables is at least 5, i.e., at least the necessary number for identifiability in a model of 6 equations. In order to apply the rank criterion for the identifiability of the saving equation (13S), say, we must consider the matrix

(14)

(I_1)	(I_2)	(I_3)	(I_4)	(H)	(E)
-1	-1	-1	-1	0	0
1	0	0	0	0	0
0	1	0	0	-1	0
0	0	1	0	0	0
0	0	0	1	0	-1

There are several ways in which a non-vanishing determinant of order 5 can be selected from this matrix. One particular way is to take the columns labeled I_1, I_2, I_3, H, E. It follows that if the present model is valid, the savings equation is indeed identifiable.

It is easily seen that the same conclusion applies to the equations explaining investment decisions of the types I_1 and I_3. Let us now inspect the rank criterion

* We have again added a term with Y_{-1} on grounds similar to those stated with respect to ($13I_2$).

matrix for the identifiability of $(13\,I_2)$:

(15)
$$\begin{bmatrix} (S) & (I_1) & (I_3) & (I_4) & (E) \\ 1 & -1 & -1 & -1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Again the determinant value of this square matrix of order 5 is different from zero. Hence the housing equation is identifiable. A similar analysis leads to the same conclusion regarding the equation $(13\,I_4)$ for quasi-investment.

Taking a look at the way in which these conclusions were reached, one notices that identifiability was attained not through the mere subdivision of total investment, but as a result of the introduction of specific explanatory variables applicable to some but not all components of investment. Whenever such specific variables are available in sufficient number and variety of occurence, on good grounds of economic theory as defined above, the door has been opened in principle to statistical inference regarding behavior parameters - inference conditional upon the assumptions derived from "theory".

How wide the door has been opened, i.e., how much accuracy of estimation can be attained from given data, is of course a matter depending on many circumstances, and to be explored separately by the appropriate procedures of statistical inference.* In the present case, the extent to which the exclusion of H and/or E from certain equations contributes to the reliability of estimates of their parameters depends very much on whether or not there are pronounced differences in the time-paths of the three predetermined variables Y_{-1}, H, E, i.e., the variables determined either exogenously or in earlier time units. These time-paths represent in a way the basic patterns of movement in the economic model considered, such that the time-paths of all other variables are linear combinations of these three paths, modified by disturbances. If the three basic paths are sufficiently distinct, conditions are favorable for estimation of identifiable parameters. If there is considerable similarity between any two of them, or even if there is only a considerable multiple correlation between the three, conditions **are adverse.**

6. Implications of the choice of the model. It has already been stressed repeatedly

* We are not concerned here with an evaluation of the particular estimation procedures applied by Ezekiel.

that any statistical inference regarding identifiable parameters of economic behavior is conditional upon the validity of the model. This throws great weight on a correct choice of the model. We shall not attempt to make more than a few tentative remarks about the considerations governing this choice.*

It is an important question to what extent certain aspects of a model of the kind considered above are themselves subject to statistical test. For instance, in the model (13) we have specified linearity of each equation, independence of disturbances in successive time units, time lags which are an integral multiple of the chosen unit of time, as well as exclusions of specific variables from specific equations. It is often possible to subject one particular aspect or set of specifications of the model to a statistical test which is conditional upon the validity of the remaining specifications. This is, for instance, the case with respect to the exclusion of any variable from any equation whenever the equation involved is identifiable even without that exclusion. However, at least three difficulties arise which point to the need for further fundamental research on the principles of statistical inference.

In the first place, on a given basis of maintained hypotheses (not subjected to test) there may be several alternative hypotheses to be tested. For instance, if there are two variables whose exclusion, either jointly or individually, from a given equation is not essential to its identifiability, it is possible to test separately (a) the exclusion of the first variable, or (b) of the second variable, or (c) of both variables simultaneously, as against (d) the exclusion of neither variable. However, instead of three separate tests, of (a) against (d), (b) against (d) and (c) against (d), we need a procedure permitting selection of one of the four alternatives (a),(b),(c),(d). An extension of current theory with regard to the testing of hypotheses, which is concerned only with choices between two alternatives, is therefore needed.

---

* In an earlier article (11) I have attempted, in a somewhat different terminology, to discuss that problem. That article needs rewriting in the light of subsequent developments in econometrics. It unnecessarily clings to the view that each structural equation represents a causal process in which one single dependent variable is determined by the action upon it of all other variables in the equation. Moreover, use of the concept of identifiability will contribute to sharper formulation and treatment of the problem of the choice of a model. However, the most serious defect of the article, in my view, cannot yet be corrected. It arises from the fact that we do not yet have a satisfactory statistical theory of choice among several alternative hypotheses.

Secondly, if certain specifications of a model can be tested given all other specifications, it is usually possible in many different ways to choose the set of "other" specifications which is not subjected to test. It may not be possible to choose the minimum set of untested specifications in any way so that strong a priori confidence in the untested specifications exists. Even in such a case, it may nevertheless happen that for any choice of the set of untested specifications, the additional specifications, while confirmed by test, also inspire some degree of a priori confidence. In such a case, the model as a whole is more firmly established than any selected minimum set of untested specifications. However, current theory of statistical inference provides no means of giving quantitative expression to such partial and indirect confirmation of anticipation by observation.

Finally, if the choice of the model is influenced by the same data from which the structural parameters are estimated, the estimated sampling variances of these estimated parameters do not have that direct relation to the reliability of the estimated parameters which they would have if the estimation were based on a model of which the validity is given a priori with certainty.

7. For what purposes is identification necessary? The question should finally be considered why it is at all desirable to postulate a structure behind the probability distribution of the variables and thus to become involved in the sometimes difficult problems of identifiability. If we regard as the main objective of scientific inquiry to make prediction possible and its reliability ascertainable, why do we need more than a knowledge of the probability distribution of the variables to permit prediction of one variable on the basis of known (or hypothetical) simultaneous or earlier values of other variables?

Knowledge of the probability distribution is in fact sufficient whenever there is no change in the structural parameters between the period of observation from which such knowledge is derived and the period to which the prediction applies. However, in most practical situations it is required to predict the values of one or more economic variables, either under changes in structure that come about independently of the economist's advice, or under hypothetical changes in structural parameters that can be brought about through policy based in part on the prediction made. In the first case knowledge may, and in the second case it is likely to, be available as to the effect of such structural change on the parameters. An example of the first case is a well-established change in consumer's preferences. An example of the second case is a change in the average level or in the progression of income tax rates.

In such cases, the "new" distribution of the variables on the basis of which predictions are to be constructed can only be derived from the "old" distribution prevailing before the structural change, if the known structural change can be applied to identifiable structural parameters, i.e. parameters of which knowledge is implied in a knowledge of the "old" distribution combined with the a priori considerations that have entered into the model.*

* For a fuller statement of the relation of the identification problem to that of prediction after structural change see Hurwicz (7).

REFERENCES

1. J.B.D. Derksen, "Long cycles in residential building: an explanation." Econometrica, April 1940, pp. 97-116.

2. M. Ezekiel, "Saving, consumption and investment." American Economic Review, March 1942, pp. 22-49; June 1942, pp. 272-307.

3. M. Ezekiel, "The statistical determination of the investment schedule." Econometrica, January 1944, pp. 89-90.

4. R. Frisch, "Pitfalls in the statistical construction of demand and supply curves," Veröffentlichungen der Frankfurter Gesellschaft für Konjunkturforschung, Neue Folge, Heft 5, Leipzig, 1933.

5. R. Frisch, "Statistical versus theoretical relations in economic macrodynamics," Mimeographed document prepared for a League of Nations conference concerning Tinbergen's work, 1938.

6. T. Haavelmo, "The probability approach in econometrics," Econometrica, Vol. 12, Supplement, 1944, also Cowles Commission Paper, New Series, No.4.

7. L. Hurwicz, "Generalization of the concept of identification." in Statistical Inference in Dynamic Economic Models, Cowles Commission Monograph 10, forthcoming.

8. L. Klein, "Pitfalls in the statistical determination of the investment schedule." Econometrica, July-October 1943, pp. 246-258.

9. L. Klein, "The statistical determination of the investment schedule: a reply." Econometrica, January 1944, pp. 91-92.

10. T.C. Koopmans, "The logic of econometric business cycle research," Journal of Political Economy, Vol. XLIX, 1941, pp. 157-181.

11. T.C. Koopmans, "Measurement without theory," The Review of Economic Statistics, Vol. XXIX, No.3, August 1947, pp. 161-172, also Cowles Commission Paper, New Series, No. 25.

12. T.C. Koopmans, G. Rasch and O. Reiersøl, "Identification as a problem in inference," to be published.

13. T.C. Koopmans, H. Rubin and R.B. Leipnik, "Measuring the equation systems of dynamic economics," in Statistical Inference in Dynamic Economic Models, Cowles Commission Monograph 10, forthcoming.

14. J. Marschak, "Economic interdependence and statistical analysis," in Studies in Mathematical Economics and Econometrics, in memory of Henry Schultz, Chicago, 1942, pp. 135-150.

15. A.C. Pigou, "A method of determining the numerical values of elasticities of demand," Economic Journal, Vol. 20, 1910, pp. 636-640, reprinted as Appendix II in Economics of Welfare.

16. Henry Schultz, Theory and Measurement of Demand, Chicago, 1938.

17. A. Wald, "Note on the identification of economic relations," in Statistical Inference in Dynamic Economic Models, Cowles Commission Monograph 10, forthcoming.