The Statistical Implications of a System of Simultaneous Equations

Author(s): Trygve Haavelmo

Source: *Econometrica*, Jan., 1943, Vol. 11, No. 1 (Jan., 1943), pp. 1-12
Published by: The Econometric Society
Stable URL: https://www.jstor.org/stable/1905714

ECONOMETRICA

VOLUME 11 JANUARY, 1943 NUMBER 1

THE STATISTICAL IMPLICATIONS OF A SYSTEM OF SIMULTANEOUS EQUATIONS

By TRYGVE HAAVELMO

1. INTRODUCTION

Measurement of parameters occurring in theoretical equation systems is one of the most important problems of econometrics. If our equations were exact in the *observable* economic variables involved, this problem would not be one of statistics, but a purely mathematical one of solving a certain system of "observational" equations, having the parameters in question as unknowns. This might itself present complicated and interesting problems, such as the problem of whether or not there is a one-to-one correspondence between each system of values of the parameters and the corresponding set of all values of the variables satisfying the equation system. For example, if we have, simultaneously, a demand curve and a supply curve, the set of possible observations might be just one single intersection point, and knowing that only would not, in general, permit us to draw any inference regarding the slope of either curve.

Real statistical problems arise if the equations in question contain certain stochastical elements ("unexplained residuals"), in addition to the variables that are given or directly observable. And some such element must, in fact, be present in any equation which shall be applicable to actual observations (unless the equation in question is a trivial identity). In other words, if we consider a set of related economic variables, it is, in general, not possible to express any one of the variables as an exact function of the other variables only. There will be an "unexplained rest," and, for statistical purposes, certain stochastical properties must be ascribed to this rest, a priori. Personally I think that economic theorists have, in general, paid too little attention to such stochastical formulation of economic theories. For the necessity of introducing "error terms" in economic relations is not merely a result of statistical errors of measurement. It is as much a result of the very nature of economic behavior, its dependence upon an enormous number of factors, as compared with those which we can account for, explicitly, in our theories. We need a stochastical formulation to make simplified relations elastic enough for applications.

This is, perhaps, generally realized among econometricians. But they frequently fail to consider, in full, the *statistical* implications of assuming a system of such stochastical equations to be *simultaneously fulfilled* by the data. More specifically, if one assumes that the economic variables considered satisfy, simultaneously, several stochastic relations, it is usually *not* a satisfactory method to try to determine each of the equations separately from the data, without regard to the restrictions which the *other* equations might impose upon the same variables. That this is so is almost self-evident, for in order to prescribe a meaningful method of fitting an equation to the data, it is necessary to define the stochastical properties of *all* the variables involved (e.g., that some of them are given time series, or remain constant, etc.). Otherwise, we shall not know the meaning of the statistical results obtained. Furthermore, the stochastical properties ascribed to the variables in *one* of the equations should, naturally, not contradict those that are implied by the *other* equations. For example, suppose that X and Y are two variables satisfying the following two stochastical relations

$$Y = aX + \epsilon_1, \tag{1.1}$$

$$X = bY + \epsilon_2, \tag{1.2}$$

where ϵ_1 and ϵ_2 are assumed to be two normally and independently distributed random variables with zero means and variances equal to σ_1^2 and σ_2^2 respectively. Then, from the first equation alone, one might perhaps think that the expected value of Y, *given* X, should be equal to aX [because $E(\epsilon_1)=0$]. But if ϵ_1 and ϵ_2 are independent, as assumed, the statement $E(Y|X)=aX$ leads to a contradiction. That is seen as follows: X and Y are linear functions of ϵ_1 and ϵ_2, viz.,

$$Y = \frac{\epsilon_1 + a\epsilon_2}{1 - ab}, \tag{1.3}$$

$$X = \frac{b\epsilon_1 + \epsilon_2}{1 - ab}. \tag{1.4}$$

Therefore, X and Y are jointly normally distributed with means equal to zero. If ρ_{XY} be the correlation coefficient between X and Y, and σ_X^2, σ_Y^2 their variances, one obtains

$$E(Y|X) = \frac{\sigma_Y}{\sigma_X}\rho_{XY}X = \frac{b\sigma_1^2 + a\sigma_2^2}{b^2\sigma_1^2 + \sigma_2^2}X, \tag{1.5}$$

the right-hand side of which is, in general, different from aX.

The example above suggests the following rule for clarifying the joint distribution properties (of the observational variables), which a system of stochastical equations implies: In a system [such as (1.1)–(1.2)] of equations, involving certain specified (but not observable) stochastical variables (such as ϵ_1 and ϵ_2), the *observable* variables involved (such as X and Y) may be considered as transformations of the specified stochastical ones (the ϵ's). Therefore, the specification of the distribution of these theoretical variables (the ϵ's) permits us, usually, to calculate the joint distribution of the observable variables, or certain properties of this distribution. This *joint* distribution should be studied to clarify the stochastical relationship, which the equation system implies with respect to the observable variables.

Instead of going further into this problem in general, it will, I think, be more explicit and more suggestive for further work in this field, to argue along a particular, simple example.

2. A MODEL AND ITS MEANING

Let us consider the following simplified *tableau économique*, assuming all prices constant:

1. The propensity to consume. Assume that if the group of all consumers in the society were repeatedly furnished with the total income, or purchasing power, r per year, they would, on the average or "normally," spend a total amount $\bar{u}$ for consumption per year, equal to

$$\bar{u} = \alpha r + \beta, \tag{2.1}$$

where α and β are constants. The amount, u, *actually* spent each year might differ from $\bar{u}$. Suppose that the amount actually spent would be given by

$$u = \alpha r + \beta + x, \tag{2.2}$$

where x is a certain random residual with mean value = zero, irrespective of the value given to r.

2. The propensity to invest. Assume that, if the group of all (private) investors in the society (the fact that they are also consumers being immaterial) were repeatedly confronted with an increase, δ, over last year, in the demand for consumption goods, they would, on the average, invest an amount, $\bar{v}$, given by

$$\bar{v} = \kappa \cdot \delta \tag{2.3}$$

(the "acceleration principle"), where κ is a constant. The amount, v, actually invested each year, for a given δ, might differ from $\bar{v}$. Suppose that the amount, v, actually invested each year would be given by

$$v = \kappa\delta + y, \tag{2.4}$$

where y is a certain random residual with mean value = zero, irrespective of the value ascribed to δ.

It is on purpose that we have used different letters u and δ for "consumption" in the two schemes above. For in the process of constructing each of the two schemes one does not have in mind the fact that the variables in one of them are restricted, simultaneously, by the other. Each of the two schemes represent a hypothetical splitting of the real economic world into two separate spheres of action, the consumer action and the producer action. It would require controlled experiments to bring about the real counterparts to one or the other of the two schemes. Actually, in a closed economy, each year's increase, δ, in the demand for consumption goods, would be the result, $u_t - u_{t-1}$, of consumers' action, and this action would, in turn, be dependent on the amount of investment, because of the "closed market" identity, $r_t = u_t + v_t$.

Now, if we assume that: 1. whatever be the value of r, u is given by (2.2), and 2. whatever be the value of δ, v is given by (2.4); then the two equations will also hold if we impose $\delta = \delta_t = u_t - u_{t-1}$, and $r_t = u_t + v_t$.

Our two theoretical schemes (2.2) and (2.4), taken together, thus lead us to believe that the data for consumption u and investment v (or their sum r) which the market mechanism actually produces will satisfy the joint system

$$u_t = \alpha r_t + \beta + x_t, \tag{2.5}$$

$$v_t = \kappa(u_t - u_{t-1}) + y_t, \tag{2.6}$$

$$r_t = u_t + v_t, \tag{2.7}$$

where x_t and y_t are certain residuals as explained above. (The actual data which we consider as the real counterparts to the theoretical variables might deviate from these theoretical variables by reason of purely statistical errors of measurements. But this we shall neglect in the analysis which follows. It is not essential for our arguments.)

The economic meaning of the joint system (2.5)–(2.7) might then be described as follows: During the year t, whatever be the "would-be" normal consumption $\alpha r_t + \beta$, the individuals choose to add a certain (positive or negative) amount, x_t, to this normal consumption, depending on a complex of factors not explicitly introduced in the theory. And similarly for the investment relation (2.6).

Now, in order that the system above be of any use for the purpose of making forecasts or some other type of inference, we need some information about two things, namely, 1. the numerical values of α, β, κ, and 2. some properties of the residual variables x and y. Usually, pure economic theory does not say anything about these things. And as long as that is the case, the equations (2.5)–(2.7) do not impose any restriction whatsoever upon the observed series of u and v (or r). For we might choose α, β, κ, at pleasure, and take (2.5) and (2.6) as definitions of the residuals x and y, since these variables are not observable anyhow. To talk about estimation of α, β, and κ without ascribing, a priori, certain properties to x and y has, therefore, no sense. The theory must first be given a statistical formulation.

3. STATISTICAL FORMULATION OF THE THEORY, AND THE PROBLEMS OF ESTIMATION

Instead of trying to give a precise statistical formulation of their economic models, many economists seem to prefer a procedure as follows: They assume, taking our example, that the observed series for u_t, v_t, and r_t satisfy the relations $u_t = \alpha r_t + \beta$, $v = \kappa(u_t - u_{t-1})$, "apart from some small deviations, x_t and y_t, which vary up and down around zero." Then they proceed to "estimate" α, β, κ, from the time series observed, first, by minimizing $\sum_t (u_t - \alpha r_t - \beta)^2$ with respect to α and β, next, by minimizing $\sum_t [v_t - \kappa(u_t - u_{t-1})]^2$ with respect to κ. [This procedure we shall call, for short, "the least-squares method applied to the equations (2.5) and (2.6) separately."]

Without further specification of the model, this procedure has no foundation, and that for two main reasons. First, the notion that one can operate with some vague idea about "small errors" without introducing the concepts of stochastical variables and probability distributions, is, I think, based upon an illusion. For, since the errors are not just constants, one has to introduce some more complex notion of "small" or "large" than just the numerical values of the individual errors. Since it is usually agreed that the errors are "on the whole small" when individual errors are large only on *rare* occasions, we are led to consider, not only the size of each individual error, but also the *frequency* with which errors of a certain size occur. And so forth. If one really tries to dig down to a clear, objective formulation of the notions of "small irregular errors," or the like, one will discover, I think, that we have, at least for the time being, no other practical instrument for such a formulation than those of random variables and probability distributions. Nor is there any loss of generality involved in the application of these analytical instruments, for any variable may be "probabilized," provided we allow sufficiently complicated distribution functions.

Next, many economists, after admitting that one must introduce the notions of probability and random variables, think that the statistical specification job is done when they have stated that: For each point of time, t, x_t is a certain random variable with a certain probability distribution, and similarly for y_t. Usually it is assumed that the variables x and y have the same distribution for all values of t. To make the statement more precise: Suppose we consider N different points of time, $t=1, 2, \cdots, N$, assuming that, for $t=0$, u_0 is a given constant. Then we consider $2N$ random variables, $x_1, x_2, \cdots, x_N, y_1, y_2, \cdots, y_N$. The specification mentioned above amounts to specifying only the *marginal* probability law for each of the $2N$ random variables. The usual specification of these probability laws is that each of the N x's has the same distribution, p_x say, with mean value = zero and finite variance, σ_x^2, and that each of the N y's has the same distribution, p_y say, with mean value = zero and finite variance σ_y^2. Some people go further, assuming, e.g., that p_x and p_y are normal distributions, to be "quite sure" that the least-squares method applied to the two equations (2.5) and (2.6) separately as indicated above, will be all right. In fact, the specification described is far from any justification for this least-squares procedure. And it is easy to see why that is so.

Let ξ_{1t} and ξ_{2t}, $t=1, 2, \cdots, N$, be $2N$ observable variables connected by a relation $\xi_{1t}=k_1\xi_{2t}+k_0+\eta_t$, where k_1 and k_0 are constants, and where η_t, $t=1, 2, \cdots, N$, are N nonobservable random variables. An essential condition for using the numbers k_1^* and k_0^* which minimizes $\sum_{t=1}^{N}(\xi_{1t}-k_1^*\xi_{2t}-k_0^*)^2$ as estimates of k_1 and k_0 (although this condition alone does not insure "best" estimates) is that

$$E(\xi_{1t}|\,\xi_{2t}) = k_1\xi_{2t} + k_0, \qquad t = 1, 2, \cdots, N.$$

Therefore, in order that the least-squares procedure applied to (2.5) and (2.6) separately should be justified, we should have to have

$$E(u_t|\,r_t) = \alpha r_t + \beta, \qquad E(v_t|\,u_t - u_{t-1}) = \kappa(u_t - u_{t-1}). \tag{3.1}$$

But if we impose this, we impose at the same time a restriction upon the *joint* distribution of the x's and the y's. More explicitly, let $p(x_t, y_t)$ be the joint probability law of x_t and y_t, and consider, for a given u_{t-1}, the distribution, $p_1(u_t, v_t)$ say, of the two variables u_t and v_t. The transformations (2.5)–(2.6) give at once

$$p_1(u_t, v_t) = [1-(1+\kappa)\alpha]p\{[u_t-\alpha(u_t+v_t)-\beta], [v_t-\kappa(u_t-u_{t-1})]\} \tag{3.2}$$

as the relation between p and p_1. A restriction, such as (3.1), upon p_1 is then, evidently, also a restriction upon p. Specification of the marginal distributions of x_t and y_t such that $E(x_t)=0$, $E(y_t)=0$, does not insure the relationships (3.1), because it does not specify $p(x_t, y_t)$, hence it does not specify $p_1(u_t, v_t)$, and, therefore, it does not suffice to draw the conclusions (3.1).

Suppose, for example, that x_t and y_t are jointly normally distributed, independent, and such that $E(x_t)=E(y_t)=0$. Then, since the transformations (2.5) and (2.6) are linear, we know at once that u_t and r_t,

u_{t-1} being given, must be jointly normally distributed, and we know also that, for example, $E(u_t|r_t)$ must be a linear function, say Ar_t+B, of r_t. What is this linear function? If ρ be the correlation coefficient of u_t and r_t, and σ_u^2, σ_r^2 their variances, then the constant A, for example, is given by

$$A = \frac{\sigma_u}{\sigma_r}\rho. \tag{3.3}$$

From (2.5) and (2.6) one can easily calculate σ_u^2, σ_r^2, and ρ directly. One obtains

$$\sigma_u^2 = \frac{\sigma_x^2 + \alpha^2\sigma_y^2}{[1-(1+\kappa)\alpha]^2}, \tag{3.4}$$

$$\sigma_r^2 = \frac{(1+\kappa)^2\sigma_x^2 + \sigma_y^2}{[1-(1+\kappa)\alpha]^2}, \tag{3.5}$$

$$\rho = \frac{(1+\kappa)\sigma_x^2 + \alpha\sigma_y^2}{\sqrt{[(1+\kappa)^2\sigma_x^2 + \sigma_y^2][\sigma_x^2 + \alpha^2\sigma_y^2]}}. \tag{3.6}$$

From (3.4), (3.5), and (3.6) one gets

$$A = \frac{(1+\kappa)\sigma_x^2 + \alpha\sigma_y^2}{(1+\kappa)^2\sigma_x^2 + \sigma_y^2}. \tag{3.7}$$

If, therefore, the first equation in (3.1) should be true under these assumptions about the distribution of x_t and y_t, we should have to have $A \equiv \alpha$, for all values of the parameters involved, and that is, evidently, not true. Thus, the assumption we have made about $p(x_t, y_t)$ and the assumptions (3.1) would, in general, be inconsistent.

The essential lesson of these considerations is this: When one has to do with a system, such as (2.5)–(2.7), this system should, for statistical purposes, be considered as a system of transformations, by which to derive the joint probability distribution of the *observable* variables from the specified distribution of the error terms. And then, to avoid inconsistencies, like the ones indicated above, all formulae for estimating the parameters involved should be derived on the basis of this joint probability law of all the observable variables involved in the system. (This, I think, is obvious to statisticians, but it is overlooked in the work of most economists who construct dynamic models to be fitted to economic data.)

The correct procedure for estimating α, β, κ, in our system (2.5)–(2.7) can be formulated precisely as follows: Let $x_1, x_2, \cdots, x_N, y_1, y_2, \cdots, y_N$ be $2N$ random variables whose joint probability law, say

$$f(x_1, x_2, \cdots, x_N, y_1, y_2, \cdots, y_N), \tag{3.8}$$

is specified, except, perhaps, for a certain number of unknown parameters (such as σ_x or σ_y above). And assume that u_0 is a given constant (initial condition). Then the joint probability law, say f^*, of the $2N$ observable random variables $u_1, u_2, \cdots, u_N, r_1, r_2, \cdots, r_N$, is given by

$$
\begin{aligned}
(3.9)\qquad & f^*(u_1, u_2, \cdots, u_N, r_1, r_2, \cdots, r_N) \\
& \qquad = [1 - (1 + \kappa)\alpha]^N f(x_1, x_2, \cdots, x_N, y_1, y_2, \cdots, y_N).
\end{aligned}
$$

The problem of estimating α, β, κ, is, therefore, the problem of estimating the parameters of this joint probability law, on the basis of a sample point $(u_1, r_1, u_2, r_2, \cdots, u_N, r_N)$. The method of estimation will depend on the form of f.

If f is such that the method of maximum likelihood applied to (3.9) can be shown to give good (if not the "best") estimates of the parameters, these estimates have a great technical advantage besides, namely that of being invariant for transformations of the parameters. That is, a function of the maximum-likelihood estimates of α, β, κ, as derived from (3.9) is equal to the maximum-likelihood estimate of that function. In particular, if we have the maximum-likelihood estimates, α°, β°, κ°, say,[1] of α, β, κ, we have also the maximum-likelihood estimates of functions of α, β, κ, appearing as coefficients in equations derived by combining (2.5), (2.6), and (2.7), simply by substitution of α°, β°, κ°, in these functions. For other methods of estimation such a procedure would, in general, *not* be permissible.

As an illustration suppose that, in the system (2.5)–(2.7), u_0 is a given constant, while $x_1, x_2, \cdots, x_N, y_1, y_2, \cdots, y_N$ are $2N$ normally and independently distributed random variables with zero means and variance equal to σ_x^2 for all the x's, and equal to σ_y^2 for all the y's. Then the maximum-likelihood estimates, α°, β°, κ°, $\sigma^\circ{}_x^2$, $\sigma^\circ{}_y^2$, of the parameters are given by

$$(3.10)\qquad \sum_{t=1}^{N} \{ [u_t - \alpha^\circ r_t - \beta^\circ][r_t + (1 + \kappa^\circ)(\beta^\circ - u_t)] \} = 0,$$

$$(3.11)\qquad \sum_{t=1}^{N} \{ [v_t - \kappa^\circ u_t + \kappa^\circ u_{t-1}][(1 - \alpha^\circ)(u_t - u_{t-1}) - \alpha^\circ v_t] \} = 0,$$

$$(3.12)\qquad \sum_{t=1}^{N} u_t - \alpha^\circ \sum_{t=1}^{N} r_t - N\beta^\circ = 0,$$

$$(3.13)\qquad \sum_{t=1}^{N} (u_t - \alpha^\circ r_t - \beta^\circ)^2 = N\sigma^\circ{}_x^2$$

[1] These somewhat unusual notations for maximum-likelihood estimates have been necessitated for purely typographical reasons.—MANAGING EDITOR.

$$\sum_{t=1}^{N} [v_t - \kappa^\circ(u_t - u_{t-1})]^2 = N\sigma^\circ{}_v{}^2. \tag{3.14}$$

If, for example, one wants the maximum-likelihood estimate of the constant A in (3.7), one has only to introduce the solutions α°, κ°, $\sigma^\circ{}_x{}^2$, $\sigma^\circ{}_y{}^2$, of the system above in the expression (3.7).

The formulae (3.10)–(3.12) above are particular instances of an easily established general rule, which may be stated as follows:

Consider m observable economic time series, say w_{1t}, w_{2t}, $\cdots$, w_{mt}, satisfying a system of m linear lag equations of the following type:

$$\begin{aligned}\sum_{i=0}^{\omega} a_{k1,i}w_{1,t-i} + \sum_{i=0}^{\omega} a_{k2,i}w_{2,t-i} + \cdots \\ + \sum_{i=0}^{\omega} a_{km,i}w_{m,t-i} + a_{k0,0} = \epsilon_{k,t}, \\ t = 1, 2, \cdots, N > \omega, k = 1, 2, \cdots, m,\end{aligned} \tag{3.15}$$

or, for short, denoting the left-hand side by $L_{k,t}$,

$$L_{k,t} = \epsilon_{k,t}, \qquad \begin{aligned} t &= 1, 2, \cdots, N, \\ k &= 1, 2, \cdots, m, \end{aligned} \tag{3.15'}$$

where, in each of the m equations, one of the coefficients $a_{k1,0}$, $a_{k2,0}$, $\cdots$, $a_{km,0}$ (the coefficient of the "dependent variable") is equal to -1 by definition, and where the mN random variables $\epsilon_{k,t}$ are independently and normally distributed variables with zero means and variances σ_k^2, $k=1, 2, \cdots, m$, which are independent of t. ω denotes the largest lag occurring in the system (3.15). Many of the coefficients a may, of course, be zero. All those a's which are not either equal to zero or equal to -1 by definition, are assumed to be unknown parameters to be estimated. Let the quantities $w_{j,t}$ for $t=0, -1, -2, \cdots, -\omega$, $j=1, 2, \cdots, m$ (i.e., the initial conditions) be *given constants*. Then the system (3.15) represents a linear transformation of the mN random variables $\epsilon_{k,t}$, $k=1, 2, \cdots, m$, $t=1, 2, \cdots, N$, in terms of the observable random variables w. Assume that the transformation is nonsingular, and let J be its Jacobian. J will, in general, be a function of the parameters a. The maximum-likelihood estimates, a°, of the unknown parameters a are then the solutions of a system of "normal equations" obtained by setting

$$\frac{1}{N}\,\frac{\partial|J^\circ|}{\partial \overset{\circ}{a}_{kj,i}}\cdot \sum_{t=1}^{N} L^{\circ 2}_{k,t} - |J^\circ| \cdot \sum_{t=1}^{N} L^\circ_{k,t}\,\frac{\partial L^\circ_{k,t}}{\partial \overset{\circ}{a}_{kj,i}} = 0 \tag{3.16}$$

for all those a's, and those only, which are not either equal to zero or equal to -1 by definition.

If, therefore, the Jacobian is independent of the a's, the method reduces to that of fitting each equation separately, by the least-squares method in the usual way, as explained previously. Also, if the Jacobian is independent of the coefficients of *some* of the equations, these particular equations may likewise be fitted, separately, by the ordinary least-squares method.

4. EQUATIONS OF PREDICTION VERSUS EQUATIONS OF THEORY

The economist may have two different purposes in mind when he constructs a model like (2.5)–(2.7).

First, he may consider himself in the same position as an astronomer; he cannot interfere with the actual course of events. So he sets up the system (2.5)–(2.7) as a tentative description of the economy. If he finds that it fits the past, he hopes that it will fit the future. On that basis he wants to make predictions, assuming that no one will interfere with the game.

Next, he may consider himself as having the power to change certain aspects of the economy in the future. If then the system (2.5)–(2.7) has worked in the past, he may be interested in knowing it as an aid in judging the effect of his intended future planning, because he thinks that certain elements of the old system will remain invariant. For example, he might think that consumers will continue to respond in the same way to income, no matter from what sources their income originates. Even if his future planning would change the investment relation (2.6), the consumption relation (2.5) might remain true. He might be interested in the consumption relation only, but in order to estimate it from past observations, he has to set up the investment equation (2.6) too. For the fact that this equation, by assumption, was active during the past, in addition to (2.5), has implications for the statistical material available—and, hence, for the technique to be used—for estimating the consumption relation (2.5) from the observations. This we discussed in the preceding section.

Let us first consider certain problems of prediction in our model (2.5)–(2.7). Assume that we know the true values of α, β, κ, and that, for simplicity, x_t and y_t are normally and independently distributed with zero means and variances equal to σ_x^2 and σ_y^2 respectively, and that their distribution is independent of *past* values of u, r, x, and y. Consider u_{t-1} as a given quantity, and consider the following problems of prediction:

1. To predict u_t from past observations.
2. To predict v_t from past observations.

3. To predict u_t when r_t and past observations are given.
4. To predict v_t when u_t and past observations are given.

As the best prediction of a variable we choose its expected value. Since, by assumption, u_t and v_t are jointly normally distributed, the conditional expectations in question can be obtained very easily from the transformations (2.5)–(2.6), using standard formulae associated with the normal distribution. For the four prediction formulae required above, one obtains then, in the same order:

$$E(u_t|u_{t-1}) = -\frac{\alpha\kappa}{1-(1+\kappa)\alpha}u_{t-1} + \frac{\beta}{1-(1+\kappa)\alpha}, \tag{4.1}$$

$$E(v_t|u_{t-1}) = -\frac{(1-\alpha)\kappa}{1-(1+\kappa)\alpha}u_{t-1} + \frac{\kappa\beta}{1-(1+\kappa)\alpha}, \tag{4.2}$$

$$\begin{aligned} E(u_t|r_t, u_{t-1}) &= \frac{(1+\kappa)\sigma_x^2 + \alpha\sigma_y^2}{(1+\kappa)^2\sigma_x^2 + \sigma_y^2}r_t + \frac{(1+\kappa)\kappa\sigma_x^2}{(1+\kappa)^2\sigma_x^2 + \sigma_y^2}u_{t-1} \\ &\quad + \frac{\beta\sigma_y^2}{(1+\kappa)^2\sigma_x^2 + \sigma_y^2}, \end{aligned} \tag{4.3}$$

$$\begin{aligned} E(v_t|u_t, u_{t-1}) &= \frac{\kappa\sigma_x^2 + \alpha(1-\alpha)\sigma_y^2}{\sigma_x^2 + \alpha^2\sigma_y^2}u_t - \frac{\kappa\sigma_x^2}{\sigma_x^2+\alpha^2\sigma_y^2}u_{t-1} \\ &\quad - \frac{\alpha\beta\sigma_y^2}{\sigma_x^2 + \alpha^2\sigma_y^2}. \end{aligned} \tag{4.4}$$

It is worth noticing that *none* of these formulae are the same as those obtained by omitting the error terms x and y in equations (2.5), (2.6). For prediction purposes the original equations of the system have no practical significance, they play only the role of theoretical tools by which to derive the prediction equations.

What *is* then the significance of the theoretical equations obtained by omitting the error terms in (2.5) and (2.6)? To see that, let us consider, not a problem of passive predictions, but a problem of government planning.

Assume that the Government decides, through public spending, taxation, etc., to keep income, r_t, at a given level, and that consumption u_t and private investment v_t continue to be given by (2.5) and (2.6), the only change in the system being that, instead of (2.7), we now have

$$r_t = u_t + v_t + g_t, \tag{2.7$'$}$$

where g_t is Government expenditure, so adjusted as to keep r constant, whatever be u and v, as given by (2.5) and (2.6). (2.7′) then does not impose any new restriction upon u and v, beyond that which is expressed by (2.5)–(2.6). Then, from (2.5) and (2.6) it is readily seen that

$$E(u_t | r_t) = \alpha r_t + \beta, \tag{4.5}$$

$$E(v_t | u_t - u_{t-1}) = \kappa(u_t - u_{t-1}). \tag{4.6}$$

That is, to predict consumption u_t and private investment v_t under the Government policy expressed by (2.7′) we may use the "theoretical" equations obtained from (2.5) and (2.6) by omitting the error terms x_t and y_t. This is only natural, because now the Government is, in fact, performing "experiments" of the type we had in mind when constructing each of the two equations (2.5) and (2.6).

University Institute of Economics
Oslo