Econometrica, Vol. 8, No. 4 (Oct., 1940)

THE INADEQUACY OF TESTING DYNAMIC THEORY BY COMPARING THEORETICAL SOLUTIONS AND OBSERVED CYCLES*

By TRYGVE HAAVELMO

IN MODERN business-cycle research the following proceeding is being commonly used: First, a mathematical model (a determinate dynamic system) is set up, as an attempt to describe approximately the interconnections between a set of economic variables, their time derivatives, lagged values, and so on, in terms of strict functional relations. By some statistical procedure the constants in the system are estimated from the corresponding observed time series. Then the system is "solved," i.e., the variables are expressed as explicit functions of time, involving the estimated parameters. The degree of *conformity* between these theoretical solutions and the corresponding observed time series is used as a test of the validity of the model. In particular, since most economic time series show cyclical movements, one is led to consider only mathematical models the solutions of which are cycles corresponding approximately to those appearing in the data.[1] This means that one restricts the class of admissible hypotheses by inspecting the *apparent form* of the observed time series.

This condition for a "good" theory is of course not a sufficient one, since there are in general many *different* a priori setups of theory which are capable of reproducing approximately the observed cycles. But, what is more important, it may not even be a necessary condition, and its application may result in a dangerous and misleading discrimination between theories. The whole question is connected with the *type of errors* we have to introduce as a bridge between pure theory and actual observations. Compared with actual observations, each equation in a dynamic model splits the observed variations into two parts, one part which is "explained" by the equation, and another part which is not accounted for, and which is ascribed to external factors. This kind of splitting is common to all theory. We usually consider such equations as "good" and "useful" theories if, in order to get full agreement between theory and observations, it *is and continues to be* sufficient to allow for only relatively small and random external factors.

There are two main ways in which such external factors may be

* The author feels much indebted to Dr. J. Marschak and Dr. A. Wald for reading the manuscript and for many valuable comments.

[1] E.g., M. Kalecki, ("A Macrodynamic Theory of Business Cycles," ECONOMETRICA, Vol. 3, No. 3, July, 1935, p. 336), goes so far as to impose on his system not only the condition of cyclical solutions, but also the condition of constant amplitude (no damping) in order to produce maintained oscillations which can be directly compared with the observed cycles.

combined with the theoretical setup, and they have fundamentally different consequences for our "explanation" of observed cycles.

One way is to consider the set of fundamental equations as exact equations without errors, i.e., they describe an ideal hypothetical model. We may perform certain elimination operations in order to get a set of *final* equations, each containing only one unknown time function. Solving these final equations we obtain the general theoretical form of the time expansions of each variable. Then we may adjust the parameters at our disposal in these general solutions so as to make the theoretical solutions fit the observed time series "as well as possible."[2] Following this approach we are led to consider the difference between theoretical and observed series as additive errors *superimposed* on the theoretical time movement of the variables. (This assumption underlies the method of periodogram analysis and other mechanical "decomposing" methods.) If these superimposed errors are supposed to be random, and if the observed series show clear cycles, there must be cycles in the theoretical ("error-free") time expansions. If not, we may be justified in rejecting the theory as unrealistic.

Another way is to introduce the errors explicitly in the original set of fundamental equations describing our model. Then we have to carry these errors along in the elimination process, and we end up with final equations which contain certain stochastical elements. This latter scheme turns out to be the one actually chosen in most of the modern studies of dynamic systems.[3] Usually this is not explicitly stated; but it is implied in the commonly used proceeding of fitting the fundamental equations in a dynamic system to data in order to get estimates of the coefficients. All such fittings admit some unexplained residuals—in the best case—of a random character. Such errors may not seriously affect the "explanatory" value of each fundamental equation taken separately, i.e., we may get a close connection between "calculated" and "observed" values for each of these equations. Here the unexplained residuals enter merely as errors of estimation, and they may be small. But the same does not apply to the *form* of the solution we obtain from the final stochastical equations. Here the

[2] See, e.g., F. W. Dresch, "A Simplified Economic System with Dynamic Elements," Cowles Commission for Research in Economics, *Report of Fifth Annual Research Conference* . . . , 1939, pp. 18–21. He writes (p. 20): "In such a solution each variable of the system will be expressed as a function of time in terms of the constants defining the assumed functions. By choosing the values of these constants in such a way that the theoretical time curves for these variables correspond as well as possible (in some sense or other) to the observed time series for these variables, one can "fit" such a model to the actual economy."

[3] See, in particular, several recent works of Professor Tinbergen and his followers.

error terms play a much more fundamental role. Indeed, the form of the final solutions (the time series) obtained if errors are neglected throughout and the form obtained when the errors are taken account of may show widely different patterns. From the fact that each fundamental equation taken separately shows only small random errors (i.e., our model is realistic) there does not necessarily follow any close similarity between the *form* of the observed series and that of the theoretical solutions obtained by neglecting the errors. In particular, the solutions obtained by neglecting all error terms may not show cycles at all, while clear cycles appear as soon as the error terms are included. Therefore, if we admit certain error terms in the system of fundamental equations, we have to investigate the effect of these errors upon the shape of the final solutions. It must be noticed that the assumption of errors in the system of fundamental equations is—except in very particular cases—incompatible with an assumption of simple superimposed errors in the final solutions as described above.

We shall now consider a constructed example which will throw some light upon these questions. We shall choose a very simple scheme, frequently occurring in economic dynamics.

Let $x(t)$ be an observable economic time series, and let the elimination result of a dynamic system be

$$(1)\quad x(t) + a_1 x(t-1) + a_2 x(t-2) = \epsilon(t-1) \qquad (t = t_0 + 2, t_0 + 3, \cdots),$$

where a_1 and a_2 are constants and $\epsilon(t)$ a random variable with expectation equal to zero and constant finite variance for all integral values of t and zero elsewhere; t_0 is the initial point of time. This means that $x(t)$ is not uniquely determined by the past, because each "year" new things happen. Such assumptions, in one form or another, underlie all dynamic theories which claim to have some relevance to facts. We write $\epsilon(t-1)$ instead of $\epsilon(t)$ because it is more realistic to assume that the external factors do not have immediate consequence for the variable considered. Moreover, it simplifies our example as shown below.

Now let us first consider the solution of (1). The principal solution of the *homogeneous* equation

$$(2)\qquad x(t) + a_1 x(t-1) + a_2 x(t-2) = 0$$

is

$$(3)\qquad x_1(t) = Ae^{\rho_1(t-t_0)} + Be^{\rho_2(t-t_0)},$$

where

$$(4)\qquad e^{\rho_1} = \frac{-a_1 + \sqrt{a_1^2 - 4a_2}}{2}, \qquad e^{\rho_2} = \frac{-a_1 - \sqrt{a_1^2 - 4a_2}}{2},$$

ρ_1 and ρ_2 being real or complex numbers, and A and B arbitrary constants. Now let

$$\lambda(\tau) \qquad (\tau \geqq 0) \tag{5}$$

be a particular integral of (2) obtained from (3) by assigning to A and B certain specific values to be disposed of later. $\lambda(\tau)$ therefore satisfies the difference equation

$$\lambda(\tau + 2) + a_1\lambda(\tau + 1) + a_2\lambda(\tau) = 0 \qquad (\tau \geqq 0). \tag{6}$$

Consider the expression:

$$v(t) = \sum_{0}^{t-t_0}{}_{\tau}\lambda(\tau)\cdot\epsilon(t - \tau). \tag{7}$$

We may write down the identity

$$\begin{aligned} v(t) + a_1v(t - 1) + a_2v(t - 2) \equiv & \sum_{0}^{t-t_0}{}_{\tau}\lambda(\tau)\epsilon(t - \tau) \\ & + a_1 \sum_{0}^{t-t_0-1}{}_{\tau}\lambda(\tau)\epsilon(t - \tau - 1) + a_2 \sum_{0}^{t-t_0-2}{}_{\tau}\lambda(\tau)\epsilon(t - \tau - 2). \end{aligned} \tag{8}$$

The right-hand side of (8) may be written

$$\begin{aligned} & \sum_{0}^{t-t_0-2}{}_{\tau}\lambda(\tau + 2)\epsilon(t - \tau - 2) + a_1 \sum_{0}^{t-t_0-2}{}_{\tau}\lambda(\tau + 1)\epsilon(t - \tau - 2) \\ & \qquad + a_2 \sum_{0}^{t-t_0-2}{}_{\tau}\lambda(\tau)\epsilon(t - \tau - 2) \\ & \qquad + \lambda(0)\epsilon(t) + \lambda(1)\epsilon(t - 1) + a_1\lambda(0)\epsilon(t - 1) \\ & \equiv \sum_{0}^{t-t_0-2}{}_{\tau}\{\epsilon(t - \tau - 2)[\lambda(\tau + 2) + a_1\lambda(\tau + 1) + a_2\lambda(\tau)]\} \\ & \qquad + \lambda(0)\epsilon(t) + [a_1\lambda(0) + \lambda(1)]\epsilon(t - 1), \end{aligned}$$

where the first expression in square brackets = 0 because of (6). Hence

$$v(t) + a_1v(t - 1) + a_2v(t - 2) = \lambda(0)\epsilon(t) + [a_1\lambda(0) + \lambda(1)]\epsilon(t - 1). \tag{9}$$

Now we shall dispose of the arbitrary constants in $\lambda(\tau)$ so that

$$\begin{aligned} \lambda(0) &= 0, \\ \lambda(1) &= 1. \end{aligned} \tag{10}$$

Then $v(t)$ satisfies a difference equation which is identical with (1). The complete solution of (1) is therefore, by adding (3) and (7),

(11) $$x(t) = x_1(t) + \sum_{0}^{t-t_0}{}_{\tau}\lambda(\tau)\epsilon(t - \tau).$$

$\lambda(\tau)$ is here completely determined by (6) and (10). It is seen that, if (3) represents a damped sine curve or two damped exponentials, then $x_1(t)$ will practically vanish after some time and then $x(t)$ depends almost entirely upon the cumulation process $v(t)$. Now $v(t)$, being a linear combination of ϵ's, is a random variable with expectation equal to zero for each point of time taken separately. But $v(t)$ and $v(t-1)$ will in general be serially correlated, the correlation depending upon the weights $\lambda(\tau)$. When the weights $\lambda(\tau)$ are damped the variances $\sigma^2_{v(t)}$ of $v(t)$ will be finite for every t, and for large $(t-t_0)$ they approach the upper limit

(12) $$\sigma_v^2 = \sigma_\epsilon^2 \cdot \sum_{0}^{\infty}{}_{\tau}\, \lambda^2(\tau),$$

where the sum is certainly convergent when $\lambda(\tau)$ is damped exponentially. The smaller the damping the greater will be the average amplitude of $x(t)$. It is well known that, when $\lambda(\tau)$ is a damped sine curve, the series $v(t)$ will show distinct cyclical movements with a principal period corresponding on the average to that of the harmonic factor in $\lambda(\tau)$. But also when $\lambda(\tau)$ is a sum of pure damped exponentials there will in general be some cyclical movements in $v(t)$. Intuitively this may be seen by the following simple considerations: When $\lambda(\tau)$ is damped, $v(t-\kappa)$ and $v(t)$ will, for sufficiently large κ, be practically independent in the probability sense. Therefore, as $v(t)$ has expectation zero, the average of $v(t)$ over a long period must tend to zero as the variance of $v(t)$ is finite. But because $v(t)$ is certainly not always zero, as is seen from (12), $v(t)$ will have to oscillate in some way around zero, and because of the serial correlation in $v(t)$ these oscillations will show some "stickiness." For example, positive serial correlation between $v(t)$ and $v(t-1)$ increases the probability of iterations, i.e., consecutive positive or consecutive negative terms in $v(t)$ as compared with $\epsilon(t)$.[4]

We shall discuss a constructed example corresponding to scheme (1) where the solution of the homogeneous equation (2) is a sum of two damped exponentials. The example is

(13) $$x(t) - 1.2x(t - 1) + 0.3x(t - 2) = \epsilon(t - 1).$$

The random series $\epsilon(t)$ was constructed from results of drawings in the Danish Class Lottery. Each $\epsilon(t)$ is the average of 10 independent ob-

[4] Methods for determining a priori the principal characteristics of such cumulative cycles have been worked out by Professor Frisch at the Institute of Economics, Oslo, but they have not yet been published.

servations from a rectangular distribution of the integers 0, 1, 2, · · · , 9. The expected value, 4.5, was subtracted; therefore

(14) $$E(\epsilon) = 0,$$

(15) $$\sigma_\epsilon^2 = 8.25.$$

The series $\epsilon(t)$ is shown in Figure 2.

The homogeneous equation

(16) $$x(t) - 1.2x(t-1) + 0.3x(t-2) = 0$$

has the solution

(17) $$x_1(t) = Ae^{-0.17t} + Be^{-1.04t},$$

where A and B are arbitrary constants. Choosing $\lambda(0)=0$, $\lambda(1)=1$, the weights $\lambda(\tau)$ become

(18) $$\lambda(\tau) = 2.04(e^{-0.17\tau} - e^{-1.04\tau}),$$

the graph of which is shown in Figure 1. This curve shows (apart from a constant factor) the time development of $x(t)$ which would fol-

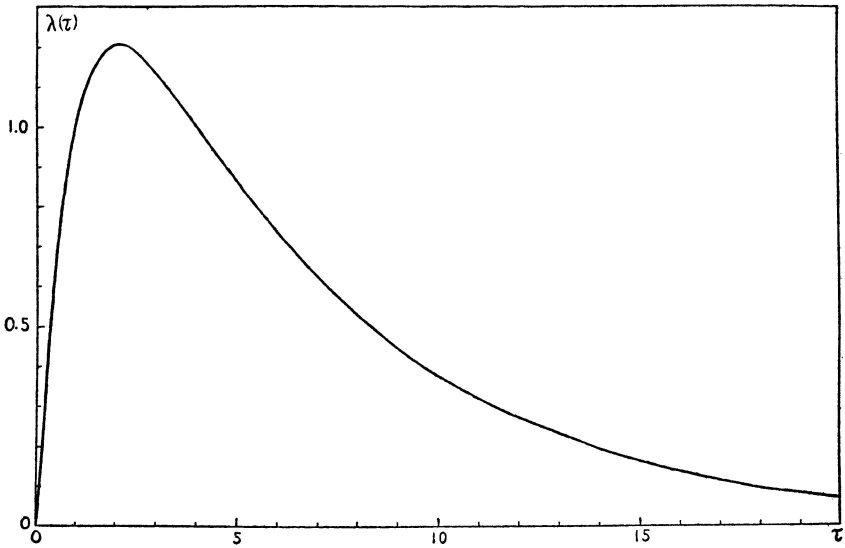

FIGURE 1.—The structural movement of the system [formula (18)].

low if $x(t)$ had been zero for two or more years, then suddenly received a positive impulse, and later were allowed to move uninterruptedly.

Now there are hardly any economic series the movements of which show resemblance to those of Figure 1: "The theory does not describe

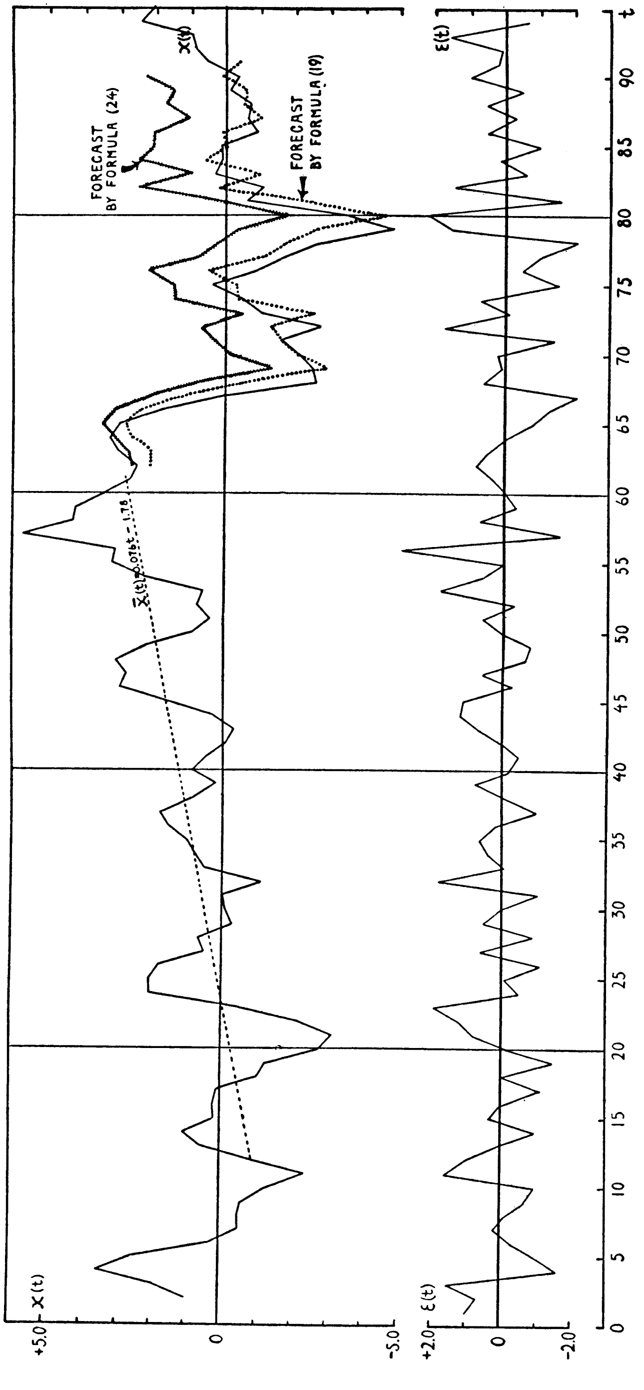

FIGURE 2.—Constructed example of shock cumulation [see formula (13)]. Lower part of figure shows the purely random series $\epsilon(t)$.

the facts." But here we must distinguish between two different ways of using the theory (16). When we want to calculate $x(t)$ in terms of the *previously observed* values $x(t-1)$ and $x(t-2)$, $\epsilon(t)$ plays only the role of errors of estimation, and if the ϵ's are small (16) is a "good" theory. But when we want to study the observed *form* of $x(t)$ we must consider the general equation (13). In Figure 2 the series $x(t)$ has been calculated step by step by means of (13), starting from two arbitrary initial conditions chosen close to zero. When we speak of cycles in economic data, we hardly think of more regular waves than the 9–10 "years" cycle shown by this curve. The series has a longer wave too, and none of these characteristic movements appear in the series λ or ϵ. The argument that a theory giving noncyclical theoretical solutions should be rejected when the observations show cycles, therefore, cannot be assumed to have general validity.[5]

Let us assume that the theory (2) has been well established on careful a priori considerations. And suppose, for example, that our observation material is the curve $x(t)$ in Figure 2 for the 52 years from $t=10$ to $t=61$, inclusive, the later observations being in the future. This part of the curve has very distinct cycles, and it also seems to have an approximately linear "trend." This "trend" looks unfavorable to our a priori theory. Nevertheless, having the general equation (1) in mind, and fitting the equation (2) to the observations, we shall obtain a useful result. Indeed, taking the regression of $x(t)$ on $x(t-1)$ and $x(t-2)$, which gives a consistent estimate of the coefficients, since $x(t-1)$ and $x(t-2)$ do not depend upon $\epsilon(t-1)$, we obtain

$$x(t) - 1.1x(t-1) + 0.26x(t-2) = 0. \tag{19}$$

A period covering 52 observations ($t=10$ to $t=61$) was used in order to have 50 observations net of each variable in (19). The solution of (19) gives two damped exponentials, $e^{-0.29t}$ and $e^{-1.05t}$, showing a little heavier damping than the theoretical. Taking the deviations between observed and calculated values of $x(t)$, we should see that they are random, and having the general equation (1) in mind, we should conclude that it is sufficient to allow for some random events in order to have full agreement between our theory and the observations. The fitting to data has added to our knowledge the fact that the structural movements of the system are exponentials, not cycles. This question was left open in the a priori theory. The theory was also *capable* of giving cycles.

[5] For actual examples from economic data where the theoretical solutions turned out to be pure damped exponentials, see the author's article "The Method of Supplementary Confluent Relations, Illustrated by a Study of Stock Prices," Econometrica Vol. 6, No. 3, July, 1938, pp. 216–218.

The usefulness of (19) for prediction purpose is shown in the dotted curve ($t=62, 63, \cdots$) in Figure 2, where $x(t)$ is forecasted *one year ahead* by means of (19), using observed values of $x(t-1)$ and $x(t-2)$. It is seen that the "best" forecast comes almost to the same as assuming that the value of x "next year" will be equal to the value "this year."

Now let us again suppose that (2) is the a priori theory, but turning to our data, i.e., $x(t)$ ($t=10, 11, \cdots, 61$), we become doubtful. There "must" be a trend, and we change our theory to take account of it. Frequently this is done even if there are no justified a priori arguments for any sort of trend. The "trend" looks fairly linear, and the cycles around it are striking. Suppose, therefore, one assumes the trend to be

$$\bar{x}(t) = kt + b \qquad (k \text{ and } b \text{ constant}). \tag{20}$$

The theory (2) is now assumed to hold in the data adjusted for trend, i.e.,

$$[x(t)-\bar{x}(t)]+a_1[x(t-1)-\bar{x}(t-1)]+a_2[x(t-2)-\bar{x}(t-2)]=0, \tag{21}$$

or, by introducing (20),

$$x(t) + a_1x(t-1) + a_2x(t-2) = c_1t + c_2, \tag{22}$$

where

$$c_1 = (1 + a_1 + a_2)k, \qquad c_2 = (1 + a_1 + a_2)b - (a_1 + 2a_2)k. \tag{23}$$

Now, fitting the equation (22) to the data $x(t)$, ($t=10, 11, \cdots, 61$) by taking the regression of $x(t)$ on $x(t-1)$, $x(t-2)$, and t ($=12, 13, 14, \cdots, 61$), we get

$$x(t) - 0.96x(t-1) + 0.37x(t-2) = 0.031t - 0.71. \tag{24}$$

The fit to the data *within* the observation period is of course here better than, or at least as good as, that obtained by (19), since we now have had two more parameters (c_1 and c_2) at our disposal. Using (23) we obtain:

$$\bar{x}(t) = 0.076t - 1.78 \qquad (t = 12, 13, \cdots, 61). \tag{25}$$

This "trend" is shown in Figure 2.

Solving the homogeneous equation (21) and inserting $a_1=-0.96$, $a_2=0.37$, we get:

$$x(t) - \bar{x}(t) = Ce^{-0.49t}\sin(m + 0.66t), \tag{26}$$

where C and m are arbitrary constants. This is a damped cycle, the period of its harmonic factor being

$$p = \frac{2\pi}{0.66} = \text{about } 9.5 \text{ ``years,''} \tag{27}$$

which is fairly close to the apparent period in the observed series. Although we still should need some external forces to outweigh the damping, since the observed cycle is not systematically damped, we now seem to have some "explanation" of the cycle. In our case we know of course from the construction of the example that this "explanation" has no sense at all. We have obtained this "explanation" by *using the apparent form of the data to formulate our theory. By doing this we have introduced into the theory and explained as structure things which are merely the effect of cumulation of random events.*

Using the equation (24) for forecast one "year" ahead in the same way as we did in the first case, we obtain the curve with cross bars ($t=62, 63, \cdots$) in Figure 2. The formula (24) is clearly useless for forecast purpose. It may be of interest to notice that the coefficients 0.37 and 0.031 in (24) are about three times their standard errors, and the coefficient 0.96 is about seven times its standard error. The coefficient of multiple correlation between $x(t)$ and the other variables is 0.88.

CONCLUSIONS

It seems possible to explain observed cyclical movements by the combination of a structure which is noncyclic, but which contains inertial forces, and outside influences of random events. This possibility should not be excluded a priori; it should be left to be determined by fitting to data.

"Correction" of *the form of* a priori *theory* by pure inspection of the *apparent shape* of time series is a very dangerous proceeding and may lead to spurious "explanations." In particular, the fitting of apparent trends which are not strongly justified on a priori reasons may lead to nonsensical results. Frequently such trend fittings will lead to the conclusion that there are later *changes in structure* (for example during the period $t=65$ to $t=68$ in our constructed example) when the real explanation is the disappearance of spurious elements introduced in our theory by the trend fitting.

University Institute of Economics
Oslo