INFORMATION AND CONTROL 6, 28–48 (1963)

Economic Processes Involving Feedback*

C. W. J. GRANGER

Princeton University, Econometric Research Program, 92-A Nassau Street, Princeton, New Jersey

and

Department of Economics, The University, Nottingham, England

The present paper is also Chapter 7 of *Spectral Analysis of Economic Time Series*, by C. W. J. Granger in association with Michio Hatanaka (Princeton University Press, 1963).

The only notation not self-evident is the following:

Let X_t be a nondeterministic stationary process with Cramér representation,

$$X_t = \int_{-\pi}^{\pi} e^{it\omega}\, dz(\omega)$$

where

$$E[dz(\omega)\, \overline{dz(\lambda)}] = 0, \qquad \omega \neq \lambda$$
$$= f(\omega)\, d\omega, \qquad \omega = \lambda,$$

$f(\omega)$ being the power spectrum of X_t.

The related process $X_t\{a(\omega), \varphi(\omega)\}$ is defined by

$$X_t\{a(\omega), \varphi(\omega)\} = \int_{-\pi}^{\pi} e^{it\omega} a(\omega) e^{i\Phi(\omega)}\, dz(\omega)$$

where

$$\Phi(\omega) = \varphi(\omega), \qquad 0 < \omega \leqq \pi$$
$$= 0, \qquad \omega = 0$$
$$= -\varphi(-\omega), \qquad -\pi \leqq \omega < 0,$$

and $a(\omega)$ is a real function of ω.

* This work was supported by a grant from the National Science Foundation.

I. FEEDBACK AND CROSS-SPECTRAL ANALYSIS

Economic variables are stochastic processes, for which the development of techniques for control involves a number of special problems. Whereas the direction of causality in physical systems is generally self-evident, for economic systems the direction of causality is usually highly debatable. It should be remembered that it is impossible to make "open loop" measurements or to carry out experiments with controlled inputs. There has recently been an awakening interest by econometricians into the possibility of using cross-spectral methods for analyzing the case in which one process $\{X_t\}$ is "causing" or is intrinsically leading another process $\{Y_t\}$. Such relationships can be expressed as

$$Y_t = X_t\{a(\omega), \varphi(\omega)\} + U_t \tag{1.1}$$

using the notation introduced above, U_t being a stationary process independent of X_t. Such models may be reasonable in the field of micro-economics, but in macro-economics there is often also a feedback equation, perhaps of the form

$$X_t = Y_t\{b(\omega), \theta(\omega)\} + V_t . \tag{1.2}$$

It should be fairly clear that when feedback is present in a system cross-spectral methods will be less appropriate. This may be shown more exactly by considering further the pair of equations (1.1) and (1.2). Assuming $E[U_t] = E[V_t] = E[U_t V_{t-\tau}] = 0$, all t, τ, they may be re-written

$$\begin{aligned} X_t &= X_t\{a(\omega)b(\omega), \varphi(\omega) + \theta(\omega)\} + U_t\{b(\omega), \theta(\omega)\} + V_t \\ Y_t &= Y_t\{a(\omega)b(\omega), \varphi(\omega) + \theta(\omega)\} + V_t\{a(\omega), \varphi(\omega)\} + U_t \end{aligned} \tag{1.3}$$

and if

$$P_t = \int_{-\pi}^{\pi} e^{it\omega}\, dz_p(\omega)$$

$$\begin{aligned} E[dz_p(\omega)\, \overline{dz_p(\lambda)}] &= 0, && \omega \neq \lambda \\ &= f_p(\omega), && \omega = \lambda \end{aligned}$$

for $P = X, Y, U, V$, then

$$dz_x(\omega) = [b(\omega)e^{-i\theta(\omega)}\, dz_u(\omega) + dz_v(\omega)]/A(\omega), \qquad \omega > 0$$

$$dz_y(\omega) = [a(\omega)e^{-i\varphi(\omega)}\, dz_v(\omega) + dz_u(\omega)]/A(\omega), \qquad \omega > 0$$

where

$$A(\omega) = 1 - a(\omega)b(\omega) \exp\{-i(\theta(\omega) + \varphi(\omega))\}$$
$$dz_x(-\omega) = \overline{dz_x(\omega)}, \qquad dz_y(-\omega) = \overline{dz_y(\omega)},$$

and if $C_r(\omega) = c_0(\omega) + iq(\omega)$ is the power cross-spectrum between $\{X_t\}$ and $\{Y_t\}$, then

$$B(\omega)f_x(\omega) = b^2(\omega)f_u(\omega) + f_v(\omega)$$
$$B(\omega)f_y(\omega) = a^2(\omega)f_v(\omega) + f_u(\omega)$$
$$B(\omega)c_0(\omega) = a(\omega)f_v(\omega) \cos\varphi(\omega) + b(\omega)f_u(\omega) \cos\theta(\omega)$$
$$B(\omega)q(\omega) = a(\omega)f_v(\omega) \sin\varphi(\omega) - b(\omega)f_u(\omega) \sin\theta(\omega),$$

where

$$B(\omega) = 1 + a^2(\omega)b^2(\omega) - 2a(\omega)\, b(\omega) \cos[\varphi(\omega) + \theta(\omega)].$$

Thus, the coherence is

$$C(\omega) = \frac{a^2(\omega)f_v^2(\omega) + b^2(\omega)f_u^2(\omega) + 2a(\omega)b(\omega)f_u(\omega)f_v(\omega) \cdot \cos[\varphi(\omega) + \theta(\omega)]}{[b^2(\omega)f_u(\omega) + f_v(\omega)][a^2(\omega)f_v(\omega) + f_u(\omega)]}$$

and the phase-diagram varies as

$$\psi(\omega) = \tan^{-1}\left\{\frac{b(\omega)f_u(\omega) \sin\theta(\omega) - a(\omega)f_v(\omega) \sin\varphi(\omega)}{b(\omega)f_u(\omega) \cos\theta(\omega) + a(\omega)f_v(\omega) \cos\varphi(\omega)}\right\}.$$

The coherence, of course, still measures the degree of dependence between the processes at each frequency, but, as now no one process is continually lagged to the other, the phase diagram is unlikely to provide useful information unless the feedback is weak ($b(\omega)$ small for all ω, $a(\omega)$ not small, say).

It should further be noted that as there are six unknown functions $(a(\omega), b(\omega), \varphi(\omega), \theta(\omega), f_u(\omega), f_v(\omega))$ and only four estimated functions $(f_x(\omega), f_y(\omega), c_0(\omega), q(\omega))$ it is not possible to estimate the unknown functions. Even if the processes $\{U_t\}$ and $\{V_t\}$ are assumed to be white noise, the estimation problem cannot be solved using spectral methods.

Thus, consideration of feedback suggests a variety of problems which will be considered in this paper:

(i) How to define feedback and test it if it is occurring in a system.

(ii) How to measure the feedback-lag (length of time taken for feedback to occur) and the strength of any feedback.

(iii) How to consider whether feedback is varying in importance and direction with frequency.

II. SOME PRELIMINARY RESULTS

If the $q \times 1$ vector $\mathbf{X}_t$ with[1]

$$\mathbf{X}_t' = \{X_{1t}, X_{2t}, \cdots, X_{qt}\},$$

$$E[X_{jt}] = 0, \qquad j = 1, \cdots, q,$$

is a multivariate, stationary, nondeterministic process, Zashuin (1941) has shown that it may be represented by

$$\mathbf{X}_t = \mathbf{B}(U)\boldsymbol{\varepsilon}_t \tag{2.1}$$

where $\mathbf{B}(U)$ is the matrix

$$\mathbf{B}(U) = [B_{jk}(U)]$$

$$B_{jk}(U) = \sum_{m=0}^{\infty} b_{jkm}U^m,$$

U being the shift operator defined by

$$UX_t = X_{t-1},$$

and $\boldsymbol{\varepsilon}_t$ is a $q \times 1$ multivariate white noise vector with mean zero, i.e.,

$$\boldsymbol{\varepsilon}_t' = \{\epsilon_{1t}, \epsilon_{2t}, \cdots, \epsilon_{qt}\}, \qquad E[\epsilon_{jt}] = 0, j = 1, \cdots, q;$$

$$E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t'] = \mathbf{I}_q; \qquad E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_{t+s}'] = \mathbf{0}_q, \qquad s \neq 0.$$

Here, $\mathbf{I}_q$ is the unit matrix and $\mathbf{0}_q$ the zero matrix, both of the qth order. Equation (2.1) may be called the "moving average" representation of the process.

Defining the theoretical autocovariances and power spectral functions by

$$\Gamma_{jk}(s) = E[X_{j,t+s}X_{kt}]$$

$$F_{jk}(\omega) = \sum_{s=-\infty}^{\infty} \Gamma_{jk}(s)e^{i\omega s}$$

$$\mathbf{F}(\omega) = [F_{jk}(\omega)], \qquad j, k = 1, \cdots, q,$$

[1] Matrix $\mathbf{A}'$ is the transpose of matrix $\mathbf{A}$.

it is easy to show that

$$\mathbf{F}(\omega) = \mathbf{B}(e^{i\omega})\mathbf{B}'(e^{-i\omega}).$$

If the equation in z, $|\mathbf{B}(z)| = 0$, has no roots on or within $|z| = 1$, we may invert (2.1) to get the "autoregressive" representation

$$\mathbf{A}(U)\mathbf{X}_t = \boldsymbol{\varepsilon}_t \tag{2.2}$$

where

$$\mathbf{A}(U) = \mathbf{B}^{-1}(U) = \sum_{j=0}^{\infty} \mathbf{A}_j U^j.$$

If we write (2.2) as

$$\mathbf{A}_0\mathbf{X}_t + (\text{past values of } \mathbf{X}_t) = \boldsymbol{\varepsilon}_t \tag{2.3}$$

we have

$$\begin{aligned} \mathbf{X}_t &= \mathbf{A}_0^{-1}\boldsymbol{\varepsilon}_t + (\text{past } \mathbf{X}_t) \\ &= \mathbf{B}_0\boldsymbol{\varepsilon}_t + (\text{past } \mathbf{X}_t). \end{aligned} \tag{2.4}$$

Equation (2.3) will be called the *basic representation* of the vector process $\{\mathbf{X}_t\}$ and Eq. (2.4) the *reduced-form representation*.

It is important to note that (2.3) is not a unique representation. If $\boldsymbol{\Lambda}$ is an orthogonal matrix, i.e., a square matrix having the property $\boldsymbol{\Lambda}\boldsymbol{\Lambda}' = \mathbf{I}_q$, then, if $\boldsymbol{\varepsilon}_t$ is a white noise vector, so is the vector $\mathbf{n}_t$ defined by $\mathbf{n}_t = \boldsymbol{\Lambda}\boldsymbol{\varepsilon}_t$, as

$$\begin{aligned} E[\mathbf{n}_t\mathbf{n}'_{t+s}] = \boldsymbol{\Lambda}E[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}'_{t+s}]\boldsymbol{\Lambda}' &= \boldsymbol{\Lambda}\mathbf{I}_q\boldsymbol{\Lambda}' = \mathbf{I}_q, \qquad \text{if } s = 0 \\ &= \boldsymbol{\Lambda}\mathbf{0}_q\boldsymbol{\Lambda}' = \mathbf{0}_q, \qquad \text{if } s \neq 0. \end{aligned}$$

Thus an alternative representation to (2.3) having exactly the same form is achieved by premultiplying (or postmultiplying) throughout by any orthogonal matrix, getting

$$\boldsymbol{\Lambda}\mathbf{A}_0\mathbf{X}_t + (\text{past } \mathbf{X}_t) = \mathbf{n}_t.$$

It follows that, as any matrix can be written as the product of an orthogonal matrix and a triangular matrix, we lose no generality by assuming $\mathbf{A}_0$ to be a triangular matrix in any basic representation. On the other hand, we are unable, in general, to assume $\mathbf{A}_0 = \mathbf{I}_q$ and still retain a basic representation having white noise as the residual term.

Viewing the reduced-form representation (2.4) with regard to predic-

tion possibilities, we see that if given all past values of $\mathbf{X}_t$ we can predict all of the present $\mathbf{X}_t$ apart from the terms $\mathbf{B}_0\boldsymbol{\varepsilon}_t$. The determinant of the covariance matrix of these terms,

$$\begin{aligned} V &= |\, E[(\mathbf{B}_0\boldsymbol{\varepsilon}_t)(\mathbf{B}_0\boldsymbol{\varepsilon}_t)']\,| \\ &= |\, \mathbf{B}_0\mathbf{B}_0'\,| = |\, \mathbf{B}_0\,|^2, \end{aligned} \tag{2.5}$$

is called the *total predictive variance* of the process $\{\mathbf{X}_t\}$. We note that premultiplying the basic form by any orthogonal matrix $\boldsymbol{\Lambda}$ has no effect on V as $|\,\boldsymbol{\Lambda}\,| = 1$.

Let us denote the set of q stochastic processes $\{X_{it}\}$, $i = 1, \cdots, q$, by Q and let $Q(j)$ be the set of processes $\{X_{it}\}$, $i = 1, \cdots, j-1, j+1, \cdots, q$, i.e., the set Q excluding $\{X_{jt}\}$, let $Q(j, k)$ be the set of processes Q excluding $\{X_{jt}\}$ and $\{X_{kt}\}$, and so forth.

For any one of the processes in the set Q, say $\{X_{it}\}$, we can form an optimum linear predictor using only those processes belonging to some subset J of Q by defining

$$P_{it}[J] = \sum_j \sum_{k=1}^{\infty} a_{jk} X_{j,t-k}, \qquad j \in J, \tag{2.6}$$

and choosing the coefficients a_{jk} so that the prediction error variance

$$V_i[J] = E[(X_{it} - P_{it}[J])^2] \tag{2.7}$$

is a minimum. Thus, $P_{it}[J]$ is the best (in a least-squares sense) predictor of X_{it} available to us if we restrict ourselves to using only the past values of the set of processes $\{X_{jt}\}$, $j \in J$. Clearly, the better the prediction, the lower will be the value of $V_i[J]$, and thus

$$0 < V_i[J] \leqq \sigma_i^2$$

where $\sigma_i^2 = E[X_{it}^2]$. $V_i[J] = 0$ if and only if the process X_{it} is deterministic.

III. DEFINITIONS OF CAUSALITY AND FEEDBACK

We shall initially restrict ourselves to vector stochastic processes having basic representation

$$\mathbf{A}_0\mathbf{X}_t = (\text{past } \mathbf{X}_t) + \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t$ is white noise and $\mathbf{A}_0$ is a *diagonal* matrix. As explained above, this is a restrictive assumption, but the class of such processes is impor-

tant because each process $\{X_{it}\}$, $i = 1, \cdots, q$, is now "caused" only by *past* values of X_t. The removal of this assumption is discussed below. We note that for such processes (3.1) can be written

$$\mathbf{X}_t = (\text{past } \mathbf{X}_t) + \mathbf{B}_0\boldsymbol{\varepsilon}_t,$$

where

$$\mathbf{B}_0\mathbf{B}_0' = \mathbf{V} = [V_i\delta_{ij}]$$

with

$$\delta_{ij} = 0, \qquad i \neq j$$
$$= 1, \qquad i = j.$$

Thus, V_i will be the prediction error variance of the process $\{X_{it}\}$ if all processes in the set Q are used, i.e.,

$$V_i = E[(X_{it} - P_{it}[Q])^2],$$

and the total prediction variance is given by

$$V = \prod_{i=1}^{q} V_i.$$

Using the notation of the previous section, we shall define causality of the process $\{X_{jt}\}$ by the process $\{X_{kt}\}$ within the set Q if

$$V_j[Q(k)] - V_j[Q] > 0.$$

Such a causality will be denoted by

$$\{X_{kt}\} \Rightarrow \{X_{jt}\}.$$

If

$$V_j[Q(k)] - V_j[Q] = 0,$$

there is no causality, denoted by $\{X_{kt}\} \not\Rightarrow \{X_{jt}\}$. Thus, we say that the process $\{X_{kt}\}$ is causing the process $\{X_{jt}\}$ if we are better able to predict X_{jt} using past values of X_{kt} than if we do not use these values.[2]

If we find $\{X_{kt}\} \Rightarrow \{X_{jt}\}$ and $\{X_{jt}\} \Rightarrow \{X_{kt}\}$, i.e., we have both $V_j[Q(k)] - V_j[Q] > 0$ and $V_k[Q(j)] - V_k[Q] > 0$, we say that direct feedback is present and denote this by

$$\{X_{jt}\} \Leftrightarrow \{X_{kt}\}.$$

[2] This definition of causality agrees with that proposed by Wiener (1956).

Two other types of feedback will also be distinguished:

(i) If $V_j[Q(j)] - V_j[Q] > 0$, we say that there is "internal feedback."

(ii) If causality chains are found of the type $\{X_{kt}\} \Rightarrow \{X_{it}\} \Rightarrow \{X_{jt}\} \Rightarrow \{X_{kt}\}$ we may talk of "indirect feedback."

Strictly speaking, we should always indicate the basic set Q of processes within which we are working whenever the causality or feedback notation is used. A possible notation is:

$$\{X_{jt}\} \Rightarrow \{X_{kt}\} \mid Q.$$

The reason why we must be careful on this point is because it is possible, for instance, that there exists a stochastic process $\{Y_t\}$ outside of Q such that

$$\{X_{kt}\} \Rightarrow \{Y_t\} \mid R$$

$$\{Y_t\} \Rightarrow \{X_{jt}\} \mid R$$

and

$$\{X_{kt}\} \not\Rightarrow \{X_{jt}\} \mid R$$

but

$$\{X_{kt}\} \Rightarrow \{X_{jt}\} \mid Q,$$

where R is the set Q plus the process $\{Y_t\}$. In such a case we are finding causality within Q due to the indirect causality via $\{Y_t\}$ in the larger set R. Putting this another way, causality is found in Q between X_{kt} and X_{jt} because X_{kt} contains information about the missing process Y_t which itself contains information about X_{jt}. However, whenever the basic set of processes Q within which we are working is clearly evident the generalized notation will not be used.

IV. TIME-LAGS CONNECTED WITH CAUSALITY AND FEEDBACK

Suppose that we have a set Q of stochastic processes and that a causality has been found between two of them $\{X_{kt}\} \Rightarrow \{X_{jt}\}$, so that we have optimum linear predictors $P_{jt}[Q]$, $P_{jt}[Q(k)]$ and prediction error variances $V_j[Q]$, $V_j[Q(k)]$ associated with them and having the property $V_j[Q(k)] > V_j[Q]$.

Define the k-truncated optimum linear predictor of X_{jt} as

$$P_{jt}[Q; k, \tau] = \sum_{p \in Q(k)} \sum_{i=1}^{\infty} a_{pi} X_{p,t-i} + \sum_{i=\tau}^{\infty} a_{ki} X_{k,t-i}$$

where the coefficients a_{ji} are chosen to minimize

$$V_j[Q; k, \tau] = E[(X_{jt} - P_{jt}[Q; k, \tau])^2],$$

$V_j[Q; k, \tau]$ being the minimum thus achieved.

If now, we find that

$$V_j[Q(k)] = V_j[Q; k, \tau] > V_j[Q],$$

then there is a *causality lag* of at least τ time units. $V_j[Q; k, \tau]$ will be a nondecreasing sequence as τ increases, and the least value of $\tau(\tau_0)$ such that

$$V_j[Q; k, \tau_0 - 1] = V_j[Q; k, \tau_0] < V_j[Q; k, \tau_0 + 1]$$

will be called the *integer causality lag* of the causality $\{X_{kt}\} \Rightarrow \{X_{jt}\}$.

We are saying here that if the causality $\{X_{kt}\} \Rightarrow \{X_{jt}\}$ occurs, but that we do not worsen our prediction of X_{jt} by not using any of the terms X_{kt}, $X_{k,t-1}$, $\cdots$, $X_{k,t-\tau+1}$, then the causality lag must be at least τ units.

The true causality lag may be $\tau_0 + a$ time units, where $0 \leqq a < 1$, as the discrete processes being considered may be samples from continuous stochastic processes and the causality need not occur exactly at one of the sampling points. To introduce such sophistication would not appear to be worthwhile in practice, but this point will be discussed again later when there is a possibility of a causality lag of less than one time unit occurring (see Section VII).

If the integer causality lags of $\{X_{kt}\} \Rightarrow \{X_{jt}\}$ and $\{X_{jt}\} \Rightarrow \{X_{kt}\}$ are τ_0 and τ_1 respectively, we shall call $\tau_0 + \tau_1$ the *integer feedback lag*.

V. STRENGTH OF CAUSALITY AND FEEDBACK

Using the notation of the previous two sections, we define the *strength* of the causality $\{X_{kt}\} \Rightarrow \{X_{jt}\}$ as

$$C(k, j) = 1 - \frac{V_j[Q]}{V_j[Q(k)]} \tag{5.1}$$

and the strength of the feedback $\{X_{kt}\} \Leftrightarrow \{X_{jt}\}$ is defined as

$$S(k, j) = C(k, j)C(j, k) = \left(1 - \frac{V_j[Q]}{V_j[Q(k)]}\right)\left(1 - \frac{V_k[Q]}{V_k[Q(j)]}\right). \tag{5.2}$$

These quantities have the properties

$$0 \leqq C(k, j) \leqq 1, \qquad 0 \leqq S(k, j) \leqq 1$$
$$S(k, j) = S(j, k).$$

The quantities measuring strength are chosen so that $C(k, j) = 0$ when $\{X_{kt}\} \not\Longrightarrow \{X_{jt}\}$ and $S(k, j) = 0$ when there is no feedback between the two processes. Although these measures have useful properties, it must be emphasized that they are arbitrarily-chosen measures and that various alternative measures could be proposed.

To indicate how these measures of strength or importance are related to alternative measures such as coherence and information, we consider the case when $q = 2$, i.e., the set Q contains only the two processes $\{X_{1t}\}$ and $\{X_{2t}\}$. Suppose that $f_1(\omega), f_2(\omega)$ are the power spectra of these processes and $C(\omega)$ is the coherence between them.

For the case $q = 2$, $V_1[Q(2)]$ will be the minimum prediction error variance of X_{1t} when only past values of X_{1t} are used, and $V_1[Q]$ will be the minimum prediction error variance of X_{1t} when past values of both X_{1t} and X_{2t} are used. $V_2[Q(1)]$ and $V_2[Q]$ are similarly defined, and $V = V_1[Q]V_2[Q]$ is the minimum total prediction variance. Kolmogoroff (1939) has shown that

$$\log V_1[Q(2)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log f_1(\omega)\, d\omega$$

and

$$\log V_2[Q(1)] = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log f_2(\omega)\, d\omega,$$

and Whittle (1955) has shown that

$$\log V = \log V_1[Q] + \log V_2[Q]$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \log [f_1(\omega)f_2(\omega)(1 - C(\omega))]\, d\omega.$$

Thus,

$$\log V_1[Q] + \log V_2[Q] = \log V_1[Q(2)] + \log V_2[Q(1)] - I$$

where

$$I = -\frac{1}{2\pi}\int_{-\pi}^{\pi} \log (1 - C(\omega))\, d\omega, \tag{5.3}$$

has been defined by Gel'fand and Yaglom (1959) as the average amount of information per unit of time contained in $\{X_{1t}\}$ about $\{X_{2t}\}$ and vice-versa. Substituting from (5.1) into this equation, we have

$$\log (1 - C(1, 2)) + \log (1 - C(2, 1)) = -I$$

or

$$1 + S(1, 2) - C(1, 2) - C(2, 1) = e^{-I}. \tag{5.4}$$

Thus, in the case of two variables, the important quantities of strength of feedback, information, and coherence are connected by Eqs. (5.3) and (5.4). In the case of q variables, similar equations exist if coherence is replaced by partial coherence, and a new concept of "partial information" is introduced.

VI. TESTS FOR CAUSALITY AND FEEDBACK

The previous sections of this chapter have been chiefly concerned with the theoretical aspects of causality and feedback, and we have been able to assume that we have available knowledge of all past values of the processes belonging to the set Q. In practice, of course, we will have only the past values of $\{\mathbf{X}_t\}$ over a finite time interval of N units, i.e., $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N$. Thus we are forced to use *approximate* linear predictors of the type

$$\tilde{P}_{jt}[J] = \sum_{p} \sum_{k=1}^{m_j} a_{pk} x_{p,t-k}, \qquad p \in J \tag{6.1}$$

which, for sufficiently large m_j, $j \in J$, will be an approximation to $P_{jt}[J]$ defined in Eq. (2.6). Although it is more general to put no limit on the truncation values m_j, we shall henceforth take all the m_j's to be equal, i.e., $m_j = m$, $j \in J$, and denote the resulting approximate linear predictor $\tilde{P}_{jt}[J, m]$. This will be an optimum predictor if the coefficients a_{jk} are chosen so that

$$\hat{V}_j[J, m] = \frac{1}{N - m} \sum_{t=m+1}^{N} x_{jt} - \tilde{P}_{jt}[J, m])^2$$

is minimized, $\hat{V}_j[J, m]$ being the resulting minimum value. In general, $E[\hat{V}_j[J, m]] > V_j[J]$, but

$$\lim_{m \to \infty} E[\hat{V}_j[J, m]] = V_j[J],$$

and so if we choose m sufficiently large the approximate linear optimum predictor will be a good approximation of the true optimum predictor.

If it is assumed that all the processes $\{X_{jt}\}$, $j \in Q$, are Gaussian, some results due to Whittle (1953) provide a test for causality. He proves that under the null hypothesis of no causality $\{X_{kt}\} \Rightarrow \{X_{jt}\}$, the statistic

$$\psi^2 = (N - q - m) \log_e \left[\frac{\hat{V}_j[Q(k), m]}{\hat{V}_j[Q, m]} \right]$$

is distributed as chi-squared with m degrees of freedom.[3] Thus, if $\hat{V}_j[Q(k),\ m]/\hat{V}_j[Q,\ m]$ becomes too large the null hypothesis of no causality will be rejected. If, and only if, both causalities $\{X_{kt}\} \Rightarrow \{X_{jt}\}$, $\{X_{jt}\} \Rightarrow \{X_{kt}\}$ are found, it can be said that feedback exists between the two processes.

The same test may be used to find the causality lag. Suppose that the null hypothesis $\{X_{kt}\} \not\Rightarrow \{X_{jt}\}$ has been rejected and that a pre-truncated approximate linear optimum predictor of X_{jt} is found,

$$\tilde{P}_{jt}[Q,\ m;\ k,\ \tau] = \sum_{p \in (k)} \sum_{i=1}^{m} a_{pi} x_{p,t-i} + \sum_{i=\tau}^{m} a_{ki} x_{k,t-i}\,,$$

being an approximation to the predictor introduced in section 7.4, with the coefficients minimizing

$$\hat{V}_j[Q, m; k, \tau] = \frac{1}{N - m} \sum_{t=m+1}^{N} [x_{jt} - \tilde{P}_{jt}]^2,$$

$\hat{V}_j[\ ]$ being the minimum.

The null hypothesis that the integer causality lag is at least $\tau + 1$ units is tested by forming the statistic

$$\psi_\tau^2 = (N - q - m) \log_e \left[\frac{\hat{V}_j[Q, m; k, \tau]}{\hat{V}_j[Q, m]} \right],$$

which, if the null hypothesis is true, will be distributed as chi-squared with $\tau - 1$ degrees of freedom. If the null hypothesis is rejected, the integer causality lag will be τ units or less.

How important is the assumption that $\{\mathbf{X}_t\}$ is a Gaussian vector process is uncertain. No equivalent test for nonnormal data exists at present, but the test given is likely to be appropriate asymptotically as $N \to \infty$. The question of whether economic series may be considered to be Gaussian, possibly after "decontamination," is debatable and considerably more research is required before a more definite statement can be made.

VII. REMOVING THE BASIC ASSUMPTION OF SECTION III

The previous four sections have all been based on the assumption that the vector process $\{\mathbf{X}_t\}$ can be represented by

$$\mathbf{A}_0 \mathbf{X}_t = (\text{past } \mathbf{X}_t) + \boldsymbol{\varepsilon}_t\,, \tag{7.1}$$

[3] If the simplifying assumption $m_j = m$, all j, had not been made, the test statistic should be

$$\psi^2 = (n - q - M/q) \log_e (\hat{V}[Q(k),\ m]/\hat{V}[Q,\ m])$$

where $M = \sum_{j=1}^{q} m_j$, and ψ^2 is distributed as chi-squared with m_k degrees of freedom under the null hypothesis.

where ε_t is white noise and $\mathbf{A}_0$ is a diagonal matrix. Using the terms since introduced, the assumption can be restated as assuming all causality lags occurring in the system to be of at least one time unit. For certain economic series such as monthly production data, such an assumption may be a realistic one, but it is certainly not so for all economic series. It will be the object of this section to study the effect of removing this assumption.

First, however, a test of whether or not the assumption holds for a given set of processes will be considered. Suppose that, using the notation of the previous section, an approximate optimum linear predictor of x_{jt} is found for all $j = 1, \cdots, q$, involving all the processes in the set Q, i.e.,

$$\tilde{P}_{jt}[Q, m] = \sum_{p=1}^{q} \sum_{k=1}^{m} a_{pk} x_{p,t-k}, \qquad \text{for each } j = 1, \cdots, q.$$

Now, forming the estimated prediction error series

$$\hat{\epsilon}_{jt} = x_{jt} - \tilde{P}_{jt}[Q, m] \tag{7.2}$$

for $j = 1, \cdots, q$ and $t = m + 1, \cdots, N$, we need to consider two possible sources of error in our approach:

(i) m has not been chosen sufficiently large, so that one or more of the approximate predictors $\tilde{P}_{jt}[Q, m]$, $j = 1, \cdots, q$, are not good approximations of the true optimum predictors $P_{jt}[Q]$, $j = 1, \cdots, q$;

(ii) The assumption that $\mathbf{A}_0$ is a diagonal matrix in representations such as (7.1) is untrue.

The possibility of complication (i) occurring can be investigated by testing for serial correlation in any of the series $\{\hat{\epsilon}_{jt}\}$. Various tests for serial correlation are available, the best known being that by Durbin and Watson (1950).

If these tests indicate that the m chosen is sufficiently large (or if new predictors and error series with m larger have been constructed) the basic assumption of (ii) can be investigated by testing for correlation between any two error series $\{\hat{\epsilon}_{jt}\}$, $\{\hat{\epsilon}_{kt}\}$ considered as independent samples. Any of the usual tests for correlation will suffice, but generally, since $N - m$ will be large, one of the quick and easy tests will be appropriate. If any correlation coefficient is found that is significantly nonzero, the assumption that $\mathbf{A}_0$ is diagonal will have to be rejected.

Once the assumption that $\mathbf{A}_0$ is diagonal is rejected, the problems of defining and testing for causality and feedback become not merely more difficult but perhaps impossible. This is shown by considering a simple

example. Suppose that the set Q consists of the two processes $\{X_t\}$, $\{Y_t\}$ only, and that there is a causality $\{X_t\} \Rightarrow \{Y_t\}$ but no feedback, and let the causality lag be exactly one time unit. Suppose now that instead of sampling the processes at times $t = 1, 2, \cdots$ the sampling had taken place at times $t = 1, 3, 5, \cdots$. In this case, the time lag is now twice the causality lag. Let the processes sampled at twice the original time unit be denoted by $\{X_T\}$, $\{Y_T\}$, and suppose further that the original representation of the processes $\{X_t\}$, $\{Y_t\}$ was

$$\begin{aligned} X_t &= \sum_{j=1}^{\infty} a_j X_{t-j} + \epsilon_t \\ Y_t &= \beta X_{t-1} + \sum_{j=1}^{\infty} b_j Y_{t-j} + \eta_t \end{aligned} \tag{7.3}$$

which, incidentally, has a causality lag of exactly one time unit. The representation for the new processes would be

$$\begin{aligned} X_T &= a_0 X_{T-(1/2)} + \sum_{j=1}^{\infty} a_j' X_{T-j} + \epsilon_T \\ Y_T &= \beta X_{T-(1/2)} + b_0 Y_{T-(1/2)} + \sum_{j=1}^{\infty} b_j' Y_{T-j} + \eta_T \end{aligned} \tag{7.4}$$

if one could define $X_{T-(1/2)}$, $Y_{T-(1/2)}$, but

$$X_{T-1/2} = \int_{-\pi}^{\pi} e^{it\omega} e^{-i\omega/2}\, dz(\omega)$$

and

$$e^{-i\omega/2} = \sum_{j=0}^{\infty} d_j e^{-ij\omega}$$

and so

$$X_{T-(1/2)} = \sum_{j=0}^{\infty} d_j X_{T-j},$$

and the representation becomes

$$\begin{aligned} X_T &= \sum_{j=1}^{\infty} a_j'' X_{T-j} + \epsilon_T' \\ Y_T &= \beta' X_T + \sum_{j=1}^{\infty} b_j'' Y_{T-j} + \eta_T', \end{aligned} \tag{7.5}$$

i.e., $\mathbf{A}_0$ is no longer diagonal.

These steps have several important implications concerning the problems of defining and testing for causality. For the original processes represented by (7.3), we cannot "predict" X_t any better by knowing Y_t, but this is not true for the new processes, as Y_T contains information about $X_{T-(1/2)}$ which can be used to better "predict" X_T. Similarly, if we know all past X_T and Y_T, we could better "predict" Y_T if we also knew X_T than if we did not, as X_T contains information about $X_{T-(1/2)}$. Hence, the definitions of causality and feedback using the idea of linear predictors are no longer appropriate. Due to the lack of uniqueness of the general representation (7.1) when $\mathbf{A}_0$ is not diagonal, it seems unlikely that causality can be defined or tested for when the causality lag is less than the time lag. In brief, the data are unsuitable for the problem being considered when $\mathbf{A}_0$ is not a diagonal matrix.

In such a case, the only appropriate method of analysis would seem to be that of classical model-building, and we are brought face-to-face with the usual problems of identification, simultaneous equation estimation, and interpretation. However, it must be emphasized that the preceding techniques have little in common with model-building methods, being based on a different philosophy. The techniques, given certain assumptions, are entirely general, whereas model-building attempts to use all the a priori knowledge and economic theory that is available.

In Section IX below, the possibility of causality strength and causality lag varying with frequency is considered. It is plausible that causality lag decreases with increasing frequency, and that by suitably filtering out high frequencies the resulting data will have a causality lag longer than the time unit. Clearly, no general rules can be laid down about such a method of making data suitable for causality testing, and each particular set of data will have to be dealt with in the light of available a priori knowledge or theory.

A further question that can be mentioned is whether or not instantaneous causality or feedback occurs in economic systems. It is the author's personal belief that instantaneous feedback does not occur, and so one can always, by sampling the processes sufficiently often, make all causality lags not less than the time unit. If this view is true, an implication is that much of the work on feedback and feedback control which is available in the field of electrical engineering cannot be applied to economics.

VIII. CALCULATIONS INVOLVED IN TESTING FOR FEEDBACK

The main calculations involved in testing for a causality $\{X_{kt}\} \Rightarrow \{X_{jt}\}$ are the formation of the approximate linear predictors of the type

$$\tilde{P}_{jt}[Q, m] = \sum_{p=1}^{q} \sum_{r=1}^{m} a_{pr} x_{p,t-r}$$

and of the resulting prediction error variance,

$$\hat{V}_j[Q, m] = \min \frac{1}{N - m} \sum_{t=m+1}^{N} (x_{jt} - \tilde{P}_{jt}[Q, m])^2$$

together with the corresponding quantities when Q is replaced by $Q(k)$.

The equations for a_{pk} when the prediction error variance is minimized are asymptotically the same as

$$\mathfrak{y}_j = \mathbf{M}\mathbf{a}, \tag{8.1}$$

where $\mathbf{a}$ is the $qm \times 1$ vector of the unknown coefficients, i.e.,

$$\mathbf{a}' = [\mathbf{a}_1', \mathbf{a}_2', \cdots, \mathbf{a}_q']$$

with (8.2)

$$\mathbf{a}_i' = [a_{i1}, a_{i2}, \cdots, a_{im}]$$

and $\mathfrak{y}_j$, $\mathbf{M}$ are $qm \times 1$ and $qm \times qm$ matrices of the estimated cross- and autocovariances, i.e.,

$$\mathfrak{y}_j' = [\mathfrak{y}_{j1}', \mathfrak{y}_{j2}', \cdots, \mathfrak{y}_{jq}']$$

with (8.3)

$$\mathfrak{y}_{jp}' = [\hat{\mu}_{jq}(1), \hat{\mu}_{jp}(2), \cdots, \hat{\mu}_{jp}(m)];$$

$$\mathbf{M} = \begin{bmatrix} \mathbf{M}_{11} & \mathbf{M}_{12} & \mathbf{M}_{13} & \cdots & \mathbf{M}_{1q} \\ \mathbf{M}_{21} & \mathbf{M}_{22} & \mathbf{M}_{23} & \cdots & \mathbf{M}_{2q} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot \\ \mathbf{M}_{q1} & \mathbf{M}_{q2} & \mathbf{M}_{q3} & \cdots & \mathbf{M}_{qq} \end{bmatrix} \tag{8.4}$$

where $\mathbf{M}_{rs} = [\hat{\mu}_{rs}(k - i)]$, i.e., the matrix with $\hat{\mu}_{rs}(k - i)$ in the ith row and kth column. Here $\hat{\mu}_{rs}(k - i)$ is the estimate of the cross-covariance,

$$\hat{\mu}_{rs}(\tau) = \frac{1}{N - \tau} \left[\sum_{i=\tau+1}^{N} x_{ri}\, x_{s,i-\tau} - \frac{1}{N - \tau} \sum_{i=\tau+1}^{N} x_{ri} \sum_{i=1}^{N-\tau} x_{si} \right]$$

where

$$E[\hat{\mu}_{rs}(\tau)] = \mu_{rs}(\tau) = E[X_{rt}X_{s,t-\tau}]$$

assuming $E[X_{rt}] = 0$, all r, t. (These quantities already appear in the cross-spectral computer programs and so need not be reprogrammed.)

Thus, the coefficients in $\mathbf{a}$ are found by forming

$$\mathbf{a} = \mathbf{M}^{-1}\boldsymbol{\mu}_j \tag{8.5}$$

and, once these are known, $\hat{V}[Q, m]$ is formed by

$$\hat{V}[Q, m] = \frac{1}{N - m}\sum_{t=m+1}^{N}(x_{jt} - \tilde{P}_{jt}[Q, m])^2. \tag{8.6}$$

To find $\hat{V}[Q(k), m]$, the same matrices are involved except that all components involving the parameter k are removed, e.g.,

$$\mathbf{a}' = [\mathbf{a}_1', \mathbf{a}_2', \cdots, \mathbf{a}_{k-1}', \mathbf{a}_{k+1}', \cdots, \mathbf{a}_q'], \quad \text{etc.}$$

Clearly, the calculations will invariably be sufficiently complicated that a high-speed electronic computer will need to be used.

The only decisions that must be made before computation starts are the size of q and m. The size of q will almost certainly be determined by the economic system being considered, and the size of k will almost certainly be limited by the ability of the computer being used to invert large matrices. A value for m of at least 10 or 12 is recommended, although experience may indicate that a larger value is required or that a smaller value is sufficient.

IX. CAUSALITY AND FEEDBACK VARYING WITH FREQUENCY

One of the important features of spectral and cross-spectral analysis is that it provides ways of observing how certain quantities vary with frequency. Just as it is possible for the strength of the relationship between two series (coherence) and the phase-lag to vary with frequency, so also is it possible that the strength of causality (and thus of feedback) and the causality lag to change with frequency. A simple (and highly unreal) example will perhaps help to show this. Consider two stock exchanges in some country, one of major importance (A) and the other of lesser importance (B). Clearly, B will be likely to follow all the fluctuations, both long-run and short-run, of A, and so we have $A \Rightarrow B$. However, A will be unlikely to be affected by the short-run fluctuations of B, but may be concerned by the long-run fluctuations. Thus, if a subscript L

denotes the low-frequency component and a subscript H the high frequency component, we may have

$$B_L \Rightarrow A_L$$

$$B_H \not\Rightarrow A_H .$$

Thus, in this example, feedback will only occur in the low frequency range.

A conceptually simple way of considering the possibility of causality and feedback changing with frequency is as follows:

Let $F_j[\ ]$, $j = 1, \cdots, m$, be a set of mutually exclusive filters such that if $\{X_t\}$ is a stationary process with Cramér representation,

$$X_t = \int_{-\pi}^{\pi} e^{it\omega}\, dz(\omega),$$

then

$$F_j[X_t] = \int_{-\pi}^{\pi} g_j(\omega) e^{it\omega}\, dz(\omega)$$

where $g_j(\omega)$ is the real function

$$g_j(\omega) = \frac{1}{m}, \qquad \frac{(j-1)\pi}{m} \leq \omega < \frac{j\pi}{m},$$

$$= 0 \quad \text{elsewhere}$$

Thus, applying the filter $F[\ ]$ cuts out all frequencies except in the band

$$\delta_j = \left(\frac{(j-1)\pi}{m}, \qquad \frac{j\pi}{m}\right)$$

and leaves the phase angle unaltered.

Let $F_j'[\ ]$ be a set of symmetric, moving-average filters approximating the set $F_j[\ ]$. If the new time series vector $\{\mathbf{y}_t(j)\}$ is formed from the original vector $\{\mathbf{x}_t\}$ by

$$F_j'[\mathbf{x}_t] = \mathbf{y}_t(j),$$

we can study the causality strengths and lags for the frequency band by carrying out the techniques introduced in the previous sections on the data $y_t(j)$. Similarly, other frequency bands can be studied by using $y_t(j)$ for the other j's.

Such a procedure will not be studied in any further detail here as the amount of computation required would be considerable in practice, and would be superimposed on what is already likely to be a very large amount of computation. However, if an intensive study of feedback within a certain group of economic series is required, the suggested technique should provide useful and interesting information, always remembering that all results will depend to a certain amount on the particular filters $F_j'[\ ]$ that are chosen.

Although the procedure outlined above is intrinsically reasonable, it has connected with it certain formidable theoretical problems. It is not possible to justify the technique theoretically by considering the perfect case when information of all the past and the actual filters $F_j[\ ]$ are available. This is because the resulting process $Y_t(j) = F_j[X_t]$ will have a frequency set of finite measure for which its spectrum is zero. Wherever this occurs, the process is deterministic, i.e. prediction can be made perfectly and so the definition of causality via prediction becomes of no use. In practice such problems have little consequence since no finite moving average filter $F_j'[\ ]$ can produce a process Y_t having zero spectrum over a set with finite measure unless the input process X_t in the definition of Y_t, i.e.,

$$Y_t = F_j'[X_t],$$

is already deterministic.

X. SUMMARY AND CONCLUSIONS

The main results of this paper can be summarized as follows:

(i) If feedback is present in a system of processes the coherence diagram will still provide useful information but the phase-diagram is unlikely to do so. (Section I.)

(ii) We say that, given a set of processes $\{X_{jt}\}$, $j = 1, \cdots, q$, there is causality of $\{X_{jt}\}$ by $\{X_{kt}\}$ if we can better predict X_{jt} using past values of X_{kt} than if we do not use these values. (Section III.)

(iii) If $\{X_{jt}\}$ causes $\{X_{kt}\}$ and also $\{X_{kt}\}$ causes $\{X_{jt}\}$, we say that feedback is present. (Section III.)

(iv) If causality $\{X_{kt}\} \Rightarrow \{X_{jt}\}$ is found, but we cannot better predict X_{jt} using the values $X_{k,t-1}, X_{k,t-2}, \cdots, X_{k,t-\tau}$ than if we did not use these values, we say that there is a causality lag of at least τ units. A measure of causality strength can be defined. (Sections IV and V.)

(v) If we assume all processes to be Gaussian, a test for causality is available. The test is likely to be appropriate asymptotically for non-Gaussian processes. (Section VI.)

(vi) When any causality lag is less than the time unit involved in the processes, no appropriate way of defining or testing for causality is suggested. (Section VII.)

(vii) It is possible that causality and feedback vary with frequency. (Section VII).

If the tests for causality are found to be efficient in practice, the theory proposed in this chapter should prove useful in testing many currently important economic hypotheses and theories, in model building, and in suggesting control methods for the economy.

One essential assumption involved in the techniques here proposed is that of stationarity. Several methods of removing trends in mean are available but the effects of more complicated nonstationarities are not known at present. The definitions of causality, feedback, lags, etc., are easily generalized to the nonstationary case by the use of nonstationary predictors, but the effect on the tests involved is likely to prove more difficult to determine.

Clearly, the basic assumption of this paper is that the future is caused by the past. It is possible to propose that the future is also caused by the expected future, but as the expectation must be based on past and present knowledge, the basic assumption still holds true.

More research is required into the field of feedback problems, particularly with respect to how feedback varies with frequency. Economic theory occasionally suggests that the direction of causality between two series will be different for the long-run and the short-run. In such a case, the over-all method will indicate a (spurious?) feedback. The method suggested in this paper for investigating the variation with frequency is clumsy, and it is to be hoped that a better and more direct method can be evolved, possibly as a generalization of the spectral method.

Finally, it should be emphasized that the "causality" defined in this paper is strictly only a second-moment causality. For nonnormal processes the true causality may be more complicated. However, just as "second-moment" prediction is a useful method, so is "second-moment" causality and feedback.

RECEIVED: September 24, 1962

References

DURBIN, J. AND WATSON, G. S., (1950), Testing for serial correlation in least squares regression. *Biometrika* **37,** 159–177.

GEL'FAND, I. M. AND YAGLOM, A. M., (1959), *Am. Math. Soc. Translations*, Ser. 2, **12,** 99.

KLEIN, L. R., (1953), "Textbook of Econometrics." Row, Peterson, New York.

KOLMOGOROFF, (1939), Sur l'interpretation et extrapolation des suites stationnaires. *Compt. rend. Acad. Sci. Paris* **208,** 2043–2045.

WHITTLE, P., (1953), The analysis of multiple stationary time series. *J. Roy. Statist. Soc. (B)*, **15,** 125–139.

WIENER, N., (1956), The theory of prediction. Chapter 8 of "Modern Mathematics for Engineers," Ser. 1, edited by E. F. Beckenback. McGraw Hill, New York.

ZASHUIN, V., (1941), *Doklady Akad. Nauk. SSSR* **23,** 435–437.