NBER WORKING PAPER SERIES

# MONEY AND INCOME CAUSALITY DETECTION

Cheng Hsiao*

Working Paper No. 167

COMPUTER RESEARCH CENTER FOR ECONOMICS AND MANAGEMENT SCIENCE
National Bureau of Economic Research, Inc.
575 Technology Square
Cambridge, Massachusetts 02139

March 1977

Preliminary: Not for quotation



*NBER Computer Research Center and University of California, Berkeley. Research supported in part by National Science Foundation Grant SOC75-18919 to the National Bureau of Economic Research, Inc.

TABLE OF CONTENTS

## I. INTRODUCTION

In economic modelling the existence and sense of direction of the causality are often chosen on *a priori* grounds. However, on many occasions economists fail to agree on the direction of changes, or whether feedback is occurring. For instance, the debate between the monetarists, who view money as a principal cause of channges in money income, and the critics of this view, who say money is a passive adapter to business conditions with little independent influence, has continued for decades. Whether one or the other side proclaims the truth is largely an empirical matter. A fundamental study of this problem has been done by Cagan in 1965 who, relying on disaggregated U.S. data from 1875-1960 and nonparametric methods, argues convincingly that the long-run relation between money supply and the price level is largely unidirectional. His analysis of the cyclical relations between money and income fails to yield a firm conclusion, however.

Recently Granger (1969) gave an explicit and testable definition of causality and feedback. Sims (1972) used a Hilbert space argument to show that Granger's definition was equivalent to the existence of a particular representation of the "driven" process with respect to the "driving" process and proposed a univariate regression method to test for the direction of causality. Since then, a significant amount of interest in the theory and technique of causality detection has been developed in the economic literature. However, the particular aspect of inference advocated by each individual author fails to yield a consensus on the particular problem being investigated. Take the case of money and income causality. Sims (1972) contended that there existed a unidirectional causality from money to GNP in the U.S. On the other hand, DyReys, Starleaf and Wang (1976)[1], Feige and Pearce (1974), etc., based on different methods, contended that money and income were, at most, weakly related. Neither is the pattern uniform for inter-country comparison. For instance, Wall (1974) found that money was the driving force for income in the U.K., while Williams, Goodhart and Gowland (1976) found the reverse causality to be the case.

The dispute, as we see it, stems at least in part from concern over the errors associated with the classical hypothesis testing where we may (i) wrongly reject the true (null) hypothesis (type I error), or (ii) wrongly accept the false (alternative) hypothesis (type II error). Take the simple example of two uncorrelated processes $\{x_t, y_t\}$. If both are first-order autoregressive processes with parameters $\alpha$ and $\beta$ respectively, then the variance of the estimated $k^{th}$ order cross-correlation ($\hat{r}_{xy}(k)$) is:

$$(1) \qquad \mathrm{Var}\ (\hat{r}_{xy}(k)) \simeq \frac{1}{T}\frac{(1+\alpha\beta)}{(1-\alpha\beta)}$$

where $T$ is the sample size. For white noise the corresponding result is:

$$(2) \qquad \mathrm{Var}\ (\hat{r}_{xy}(k)) \simeq \frac{1}{T}$$

Hence if $\alpha\beta$ is positive, (1) is inflated relative to (2), whereas if $\alpha\beta$ is negative, (1) is deflated. Equation (1) then shows that very large cross-correlations, all of them spurious, can be generated between two uncorrelated processes as a result of the large autocorrelations within the two processes (for economic examples, see Granger and Newbold [1974]).

An incorrect inference about the nature of time series data would then set a critical region which would either give too high a significance level and too low a power or vice versa. The concern of such a "spurious regression" phenomenon led Pierce and Haugh (1977) to propose causality detection based on cross-correlations of whitened series, which may be viewed as an attempt to set the correct significance level. The concern of "spurious independence," or the power of the test, led people to propose testing procedures based on a direct multivariate autoregressive-moving average model fitting, generalized least square method, etc. (Caines and Chan [1975], DyReyes, Starleaf and Wang [1976], Sims [1972, 1975], Wall [1974], etc.). In this paper we intend to survey and suggest the theoretical framework of the important aspects of causality detection with the purpose of conveying to the reader the essential features and the different forms in which inferences may be drawn from given data.

Section II presents the basic theorem characterizing the causality events and suggests two feedback detection methods which, like the one suggested by Pierce and Haugh (1977), are based on correlation analysis. In Section III we survey other well-known causality detection methods and try to relate and to compare them with the methods suggested in Section II. Section IV briefly reviews the theoretical controversy of the relationship between money and income and presents some empirical evidence based on the methods discussed in this paper. Conclusions are in Section V.

## II. CHARACTERIZATION OF FEEDBACK-FREE PROCESS AND ITS DETECTION

Let $\{X , Y \}$ be joint covariance stationary, purely linearly indeterministic, processes. Let $A_t$ be the given information set, including at least $\{x_t, y_t\}$, and $A_t - X_t$ be the information set apart from $X_t$. Let $\bar{A}_t = \{A_s : s<t\}$, $\bar{\bar{A}}_t = \{A_s : s \leq t\}$, and similarly define $\bar{X}_t$, $\bar{Y}_t$. $\bar{\bar{X}}_t$, $\bar{\bar{Y}}_t$. Denote by $\sigma^2 (Y|B)$ the mean square error of the minimum mean square error prediction of $Y_t$ given information set $B$. Granger's (1969) definition of causality and feedback are:[2]

Definition 1 (Causality): If $\sigma^2(Y|\bar{A}) < \sigma^2(Y|\overline{A-X})$, we say that $X$ is causing $Y$, denoted by $X_t \Rightarrow Y_t$ .

Definition 2 (Feedback): If $\sigma^2(Y|\bar{A}) < \sigma^2(Y|\overline{A-X})$ , and $\sigma^2(X|\bar{A}) < \sigma^2 (X|\overline{A-Y})$, we say that feedback is occurring, denoted by $x_t \Leftrightarrow y_t$ .

To give an operative meaning of these definitions, we write the linearly regular processes X and Y as:

$$(3) \qquad y_t = \sum_{j=0}^{\infty} a_j \xi_{t-j} + \sum_{j=0}^{\infty} b_j \eta_{t-j} ,$$

$$x_t = \sum_{j=1}^{\infty} c_j \xi_{t-j} + \sum_{j=0}^{\infty} d_j \eta_{t-j} ,$$

where $\xi$ and $\eta$ are mutually uncorrelated white noise processes with unit variance. In order to make the representation (3) unique, we take $c_o$ to be zero and $a_o$, $d_o$ positive. The following theorem is proved by Caines and Chan (1975), Sims (1972).

Theorem 1: If $Y \nRightarrow X$ relative to the information set A, the following statements are equivalent:

(i) c identically zero.

(ii) the least squares estimate $\hat{y}_{t|t}$ of $y_t$ given the observations $\bar{\bar{X}}_t$ is identical to the estimate $\hat{y}_{t|\infty}$ of $y_t$ given $\{x_t\}$, for $t = \ldots, -1, 0, 1, \ldots$.

(iii) if $X \Rightarrow Y$, there exists a unique representation of $y$ with respect to x of the form

$$y_t = \sum_{j=0}^{\infty} K_j \, x_{t-j} + \sum_{j=0}^{\infty} L_j \, \xi_{t-j},$$

where the processes x and $\xi$ are orthogonal.

Based on this theorem, various testing procedures have been proposed to detect feedback. In this section, we propose yet two more testing procedures. The methods are based on the idea of correlation between two time series rather than regression analysis. In regression theory, one variable is considered random or "dependent" and others fixed or "independent." In correlation several variables are considered and treated symmetrically. Since both X and Y are stochastic, we feel it is more natural to consider them in terms of a correlation approach rather than a regression approach. Of course, if we start with a joint normal distribution, we arrive at the same tests in either case when the two procedures are independent. The probability theory under the null hypothesis is the same. However, if the alternative is true, the distribution of the test criterion will be different.

Let $P_{t-1}$ be a unique orthogonal projection matrix of $x_t$ on $\bar{X}_t$ = $\{x_s | s<t\}$ and let $\Pi_t$ be a unique orthogonal projection matrix of $y_t$ on $\bar{\bar{X}}_t$. Define

(4) $v_t = (I - P_{t-1})\, x_t$ ,

and $u_t = (I - \Pi_t)\, y_t$ ,

(5) $\rho_{uv}(k) = E\,(u_t v_{t+k})$.

By construction $\rho_{uv}(k) = 0$ for $k \le 0$. Then

Lemma 1: $\rho_{uv}(k) = 0$ for all $k > 0$, if and only if $Y \not\Rightarrow X$.

Proof: A direct proof is obtained by noting that from theorem 1(ii) the partial covariances of $y_t$ and $x_{t+j}$, for $j \ge 1$ given $\bar{\bar{X}}_t$ are zero. Thus the if part follows immediately. To prove that lemma 1 implies all partial covariances between $y_t$ and $x_{t+j}$ given $\bar{\bar{X}}_t$ are zero, we note that $\rho_{uv}(1) = 0$ only when $u_t$ is orthogonal to $\bar{\bar{X}}_{t+1}$. By the inductive argument of the orthogonality principle, we conclude that $\rho_{uv}(k) = 0$ for all $k > 0$ states that $u_t$ is orthogonal to $\bar{X}_\infty$. Therefore lemma 1 holds.

To devise testing methods from lemma 1, we assume that the linear regular nondeterministic process (3) satisfies the stability condition. Therefore we can approximate it by a finite order autoregressive process. Under the null hypothesis of no feedback from y to x, the probability distribution of $x_t$ given $(\bar{X}_t - \bar{X}_{t-p})$ , $P\,(x_t | x_{t-1}, \ldots, x_{t-p})$, for p sufficiently large, would be approximately a white noise. The correlation between $v_{t+k}$ and $u_t$ will then have mean zero and asymptotic variances $(T-p)^{-1}$ . Thus, one way to test feedback is to:

Test 1: Regress $x_t$ on past x's, and denote the residual by $\hat{v}_t$. Regress $y_t$ on current and past x's and denote the residual by $\hat{u}_t$. Compute the cross-correlation coefficient between $\hat{u}_t$ and $\hat{v}_{t+k}$, $\hat{r}_{uv}(k)$, where

$$(6) \qquad \hat{r}_{uv}(k) = [c_{\hat{u}}(0) \cdot c_{\hat{u}}(0)]^{-\frac{1}{2}} \; c_{\hat{u}\hat{v}}(k) \; ,$$

and $c_{\hat{u}}(0)$, $c_{\hat{v}}(0)$ and $C_{\hat{u}\hat{v}}(k)$ are the sample variances and cross-covariances of $\hat{u}_t$, $\hat{v}_t$. The test statistic calculated would be

$$(7) \qquad \phi_1 = (T-p) \sum_{k=1}^{M} \hat{r}^2_{uv}(k)$$

In a large sample this statistic is approximately chi-square distributed with M degrees of freedom under the null hypothesis, because the asymptotic distribution of a finite set of lagged cross-correlations between two independent linear processes is normal (Hannan [1971, p. 230]). The proof of $\hat{r}_{uv}(k)$ having the same asymptotic distribution follows straightforwardly from that of Haugh (1976) except for the minor complication here that although $v_t$ is a white noise process, $u_t$ is not. However, the independence of u and v implies that $c_{uv}(k)$ is still of $O(T^{-\frac{1}{2}})$; thus, there is no change in the basic argument. Neither is there a reduction in the degrees of freedom for this test because the constraints placed on the residuals by the fitting process did not play a crucial role as that in the case of Box and Pierce (1970) or Chitturi (1974).

Quenoulli (1948) showed, by a sampling experiment, that when the two series are uncorrelated, the variances are approximately $(T-p)^{-1}$, even for small samples, and the bias is small.[4] On the other hand, if the alternative is true, the cross-correlation estimates based on moments are not necessarily efficient. The variance/covariance matrix of the cross-correlation functions are now complicated functions of the true autocorrelation functions and the true cross-correlation functions (Hannan [1971, Ch. 4]). The efficiency of the statistics, as test statistics, depends not only on their variance under the null hypothesis but also on their distribution when the null hypothesis is not true. Although it does not necessarily follow that a more powerful test will be obtained by using a statistic which is more efficient than the one which is less efficient, there is a strong intuitive appeal to use a statistic which is consistent and asymptotically efficient under both the null and the alternative. It is likely to be locally more powerful than others (see Kendall andStuart [1961, Ch. 25], Rao [1962], and the discussion in Section III below).

The reason one suspects Test 1 is not fully efficient is that the stochastic processes $(v_t, u_t)$ are not serially independent. Using spectral representation, we can trade our sequence of dependent random variables for a sequence of independent random variables. Thus, an approximate likelihood ratio test can be proposed for the two residual time series.

The null hypothesis in the frequency domain can be stated as:

$H_0$: $F(\omega)$ diagonal for all $\omega$,

where

(8) $$F(\omega) = \begin{bmatrix} f_u(\omega), & f_{uv}(\omega) \\ f_{vu}(\omega), & f_v(\omega) \end{bmatrix},$$

$$f_u(\omega) = \frac{1}{2\pi}\sum_{-\infty}^{\infty}\exp(-i\omega\tau)c_u(\tau),$$

$$f_v(\omega) = \frac{1}{2\pi}\sum_{-\infty}^{\infty}\exp(-i\omega\tau)c_v(\tau),$$

$$f_{uv}(\omega) = \frac{1}{2\pi}\sum_{-\infty}^{\infty}\exp(-i\omega\tau)c_{uv}(\tau) = \frac{1}{2\pi}\sum_{1}^{\infty}\exp(-i\omega\tau)c_{uv}(\tau) \text{ (by construction)}$$

$$c_u(\tau) = E[(u_t - Eu_t)(u_{c+\tau} - Eu_t)],\quad c_v(\tau) = E[(v_t - Ev_t)(v_{c+\tau} - Ev_t)],$$

$$c_{uv}(\tau) = E[(u_t - Eu_t)(v_{c+\tau} - Ev_t)].$$

If the spectra are sufficiently smooth, they can be divided into $M$ bands between $(0,\pi)$ such that for each band each spectral matrix is constant (see Wahba [1968] for a precise condition). We can then obtain an estimate of $\hat{F}(\omega)$ by first transforming $\hat{u}$ and $\hat{v}$ into $w_u(\omega_d)$ and $w_v(\omega_d)$, where

$$w_u(\omega_d) = (2\pi n)^{-\frac{1}{2}} \sum \hat{u}_t \exp(it\omega_d) ,$$

$$w_v(\omega) = (2\pi n)^{-\frac{1}{2}} \sum \hat{v}_t \exp(it\omega_d) ,$$

$d = 1, \ldots [n/2]$ , with $n = (T-2p)$ and $[\cdot]$ the largest integer not greater than the indicated number. Let $m$ satisfy $mM = [\frac{n}{2}]$. As $T\to\infty$ , we can group the $\omega_d$ into $M$ sets (excluding $\omega_\ell = 0$), the $\ell^{th}$, which we call $S_\ell$, containing $m$ adjacent $\omega_d$ nearest to $\omega_\ell$. We put

$$\hat{f}_u(\omega_\ell) = \frac{1}{m} \sum_{\omega_d \epsilon S_\ell} w_u(\omega_d)\, w_u^*(\omega_d)$$

$$\hat{f}_v(\omega_\ell) = \frac{1}{m} \sum_{\omega_d \epsilon S_\ell} w_v(\omega_d)\, w_v^*(\omega_d)$$

$$\hat{f}_{uv}(\omega_\ell) = \frac{1}{m} \sum_{\omega_d \epsilon S_\ell} w_u(\omega_d)\, w_v^*(\omega_d)$$

where the super asterisk indicates transposition combined with conjugation. Under appropriate conditions, such an estimate is consistent (see Anderson [1971, Ch. 10], Brillinger [1975, Ch. 7]).

The approximate likelihood ratio test for the null hypothesis is given by:

Test 2:

$$(9) \qquad \lambda = \prod_{\ell=1}^{M} \frac{|\hat{F}(\omega_\ell)|}{\hat{f}_u(\omega_\ell)\ \hat{f}_v(\omega_\ell)} = \prod_{\ell=1}^{M} [1 - \hat{R}_{uv}(\omega_\ell)] ,$$

where

$$(10) \qquad \hat{R}_{uv}(\omega_\ell) = |\hat{f}_{uv}(\omega_\ell)|^2/\hat{f}_u(\omega_\ell)\hat{f}_v(\omega_\ell) ,$$

which is a measure of causal coherence from y to x . Although (10) is different from the definition given by Granger (1969), it may be viewed as a direct generalization. We note that $F(\omega_\ell)$, $\ell=1, . . . M$ , is asymptotically indepently distributed as complex Wishart distribution $m^{-1}\ W_c\{m,F(\omega_\ell)\}$ under appropriate conditions.[5] Since the distribution in the complex case is identical with the real case except for having twice the sample size and twice the variable dimensions, we can show, by a similar manipulation as Anderson (1958, ch. 9) and Wahba (1968), that when $H_o$ is true

$$(11) \qquad \lambda \sim \prod_{\ell=1}^{M} Z_\ell$$

where $Z_\ell$ are M independent beta random variables with 2(m-1) and 2 degrees of freedom. Therefore

$$(12) \qquad \phi_2 = -\frac{1}{M}\log \lambda$$

has mean $\frac{1}{m-1}$ and variance $\frac{1}{M(m-1)^2}$ under the null hypothesis. Choose m (large) fixed and let $M\to\infty$ , it is asymptotically normally distributed.

If the sample size is large, this causal spectral analysis has conceptual advantages for conveying information about interdependent events. Usually it also has desirable properties for working with reasonably smoothed spectra. For instance, if no feedback is occurring, the residual spectra $f_v$ would be approximately a constant over the frequency. If feedback is occurring, then simply removing past effects of x would not, in general, make $f_v$ constant. On the other hand, if feedback does occur, the causal coherence squared (10) would give a measure of the strength of causality Y→X plotted against frequency. We can also define causal phase and gain by

$$(13) \qquad \psi(\omega_\ell) = \frac{1}{2\pi} \arctan\left[\frac{\operatorname{Imag}(f_{uv}(\omega_\ell))}{\operatorname{Re}(f_{uv}(\omega_\ell))}\right],$$

and

$$(14) \qquad G(\omega_\ell) = \left(\frac{f_v(\omega_\ell)}{f_u(\omega_\ell)}\right)^{1/2},$$

accordingly to give us an idea of the causal lead and amplification or attenuation from Y to X at each frequency component $\omega_\ell$.

## III. OTHER TESTING METHODS

In this section, we try to relate the two testing procedures suggested in the last section to other well-known testing procedures, which are based on prefiltering and regression methods and, to the extent possible, compare their relative power. All these tests are derived from various equivalent statements of theorem 1.

Test 3 (Sims test): From theorem 1(ii), Sims(1972) suggested an F test for causal detection. That is, first regress Y on past and future values of X, taking account by generalized least squares or prefiltering of the serial correlation in the residuals. If there is no feedback from Y to X, future values of X in the regression should have coefficients insignificantly different from zero, as a group.

Test 4 (Pierce and Haugh (1976), Wall(1974)): First use separate filters on X and Y to ensure that each is very nearly pre-whitened. Then use cross-correlation analysis to determine whether the residuals of the two prewhitened series are cross-correlated. If causality runs from X to Y only, all the cross-correlations between the residuals of X and lagged residuals of Y should be insignificantly different from zero. Haugh (1976) showed that under the null hypothesis these estimated residual cross correlations are asymptotically normally distributed with mean zero and variance 1/n. Thus

$$\phi_4 = n \sum_{k=1}^{M} \hat{r}_k^{\,2} \tag{15}$$

is chi-square distributed with M degrees of freedom.

Test 5: From theorem 1(iii), it is clear that one way to test feedback free is to test the independence between the stochastic regressors x and the disturbances v. Various testing methods have been suggested by Wu (1973). However, the power of these tests depend crucially upon the available instrumental variables. Therefore, except mentioning it as a possible candidate, we shall not elaborate on it here.

Test 6 (Direct test): From theorem 1(i), we can test for feedback free by fitting an impulse response function to model (1) and then test whether C=0. Such a procedure has been applied by Caines and Chan (1975), and Wall (1974), etc. on British data.

All these tests have the same distribution when the null hypothesis $H_o$ is true. It is, however, of no use to know merely what properties a critical region will have when $H_o$ holds. For in general we can find many, and often even an infinity, of sub-regions W of the sample space, all having the same significance level. The problem of testing a hypothesis is essentially one of choice between the tests which minimize the probability of a type II error, after controlling the probability of a type I error. Although all these tests are consistent under the alternative hypothesis in the sense that the power of the test, defined as the probability of rejecting the null hypothesis $H_o$ when $H_o$ is false, approaches one when sample size goes to infinity, their relative efficiencies are not identical. In particular, the power of these tests is not independent of the process characterizing the perceived driving forces. We use following simple examples to illustrate their interrelationships, and to the extent possible, compare their relative efficiency.

Consider the null hypothesis of the following simple model $H_o$:

(16) $$y_t = u_t$$

$$x_t = \alpha x_{t-1} + v_t$$

against the alternative $H_1$:

(17) $$x_t = \alpha x_{t-1} + \delta y_{t-1} + v_t \ ,$$

where $u_t$ and $v_t$ are serially and mutually independent Gaussian random variables with mean zero and variances $\sigma_u^2$ and $\sigma_v^2$. Then a direct test (test 6) of no feedback from y to x is to construct the statistic

(18) $$\phi_6 = \frac{X_t'A\, Y_{t-1}\, B^{-1} Y_{t-1} A X_t}{X_t' M X_t} \cdot (T-3)$$

where $Y_t$ and $X_t$ are (T-1)x1 vectors of $(y_2 \ldots, y_T)'$ and $(x_2 \ldots, x_T)$, and $\hat{\delta}$ is the least square estimate of the coefficient of $y_{t-1}$ (of the regression of $x_t$ on $x_{t-1}$ and $y_{t-1}$), $M = I - Z(Z'Z)^{-1}Z$, with I the (T-1) rowed identity matrix, $Z = [X_{t-1}, Y_{t-1}]$, $A = I - X_{t-1}'(X_{t-1}'X_{t-1})^{-1}X_{t-1}$,

$B = Y'_{t-1}Y_{t-1} - Y'_{t-1}X_{t-1}(X'_{t-1}X_{t-1})^{-1}X_{t-1}Y_{t-1}$. $\phi_6$ has an F distribution with 1 and T-3 degrees of freedom under the null hypothesis and a non-central F distribution with the non-centrality parameter $\delta B\delta/\sigma_v^2$ under the alternative hypothesis.

We also use basically the same statistic for test 1 and 2.[6] Since, under the null hypothesis, $x_t$ conditional on $x_{t-1}$ is a white noise, the classical test based on R in the case of test 1 is equivalent to using the statistic $\phi_6$. As far as test 2 is concerned, because conditional on $x_{t-1}$, $x_t$ and $y_{t-1}$ are white noise processes, their spectra and cross-spectra are flat, therefore M = 1 and the estimate of the spectral density matrix in this case is:

$$(19) \qquad \hat{F}(0) = \frac{1}{n} \sum_{j=1}^{\frac{n}{2}} \left\{ \begin{pmatrix} w_u(\omega_d) \\ w_v(\omega_d) \end{pmatrix} (w_u(-\omega_d), w_v(-\omega_d)) + \begin{pmatrix} w_u(-\omega_d) \\ w_v(-\omega_d) \end{pmatrix} (w_u(\omega_d), w_v(\omega_d)) \right\}$$

$$= \frac{2}{n} \sum_{j=1}^{\frac{n}{2}} \mathrm{Re} \begin{pmatrix} w_u(\omega_d) \\ w_v(\omega_d) \end{pmatrix} (w_u(-\omega_d), w_v(-\omega_d)).$$

Equation (19) is asymptotically distributed as a real Wishart distribution of dimension 2 and degrees of freedom n(=T-2) (Brillinger (1974), theorem 7.3.3). Therefore, $\lambda$ has a beta distribution with T-3(=n-1) and 1 degree of freedom. Since $\lambda = \frac{1}{1+(1/T-3)F_{1,T-3}}$, we have

$$(20) \qquad (T-3) \cdot \frac{1-\lambda}{\lambda} = F_{1,\, T-3} = \phi_6.$$

Test 3, which is the same as test 1 in this case except that the former uses the regression analysis while the latter uses the correlation analysis, uses the statistic:

(21) $$\phi_3 = \frac{Y'_{t-1}AX_tD^{-1}X_tAY_{t-1}}{Y'_{t-1}\tilde{M}Y_{t-1}/T-3}$$

$$= \{-1 + B(Y'_{t-1}AX_tD^{-1}X_tAY_{t-1})^{-1}\}^{-1}$$

$$= \phi_6 ,$$

where

$$\tilde{M} = I - \tilde{Z}(\tilde{Z}'\tilde{Z})^{-1}\tilde{Z}' ,$$

$$\tilde{Z} = (X_t , X_{t-1}) ,$$

$$D = [X'_tX_t - X'_tX_{t-1}(X'_{t-1}X_{t-1})^{-1}X'_{t-1}X_t] .$$

That is, Sims test in this case is identical to the likelihood ratio test in the parameteric model.

Although test 1, 2 and 3 and 6 use the same statistic, they are based on different conceptions about the joint distribution of y and x. The probability theory under the null hypothesis is the same whether one uses the correlation analysis or the regression analysis. The distribution of the test criterion when the null hypothesis is not true differs in the two cases (see Anderson (1958)). Thus the power function will be different even though the same test is valid in each case. Whether one prefers the former approach to the latter will be largely a matter of judicious choice and individual preference. However, when $T\to\infty$ , the results are identical in both situations. (Kendall and Stuart (1965, ch. 26)).

Test 4 in this case uses the statistic

(22) $$\phi_4 = \frac{X'_tAY_{t-1}(Y'_{t-1}Y_{t-1})^{-1}Y'_{t-1}AX_t}{X'_tA[I-Y_t(Y'_{t-1}Y_{t-1})^{-1}Y'_t]AX_t} \cdot (T-3) ,$$

Under the null hypothesis $\phi_4$ is again F distributed with 1 and T-3 degrees of freedom. However, under the alternative the test is less powerful than the testing procedures just mentioned. Since the difference of the numerator of $\phi_6$ and $\phi_4$ is

(23) $$X_t' A Y_{t-1}(Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' X_{t-1} \; D^{-1} X_{t-1}' Y_{t-1}(Y_{t-1}' Y_{t-1})^{-1} Y_{t-1}' A X_t,$$

which is positive definite, while the denominator of $\phi_4$ is equal to the denominator of $\phi_6$ plus (23). Thus, although test 4 has the advantage of eliminating possible spurious correlation, it overkills by losing the power of the test, which may partially contribute to the fact that most of the time series Pierce (1977) studied showed no sign of causation.

When $y_t$ follows a first order autoregressive process:

(24) $$y_t = \gamma y_{t-1} + u_t,$$

the direct test ($\phi_6$) is independent of such a change in the characterization of the perceived driving forces, but not for other testing methods. For instance, to apply Sims test without taking account the serial correlation in the residual is identical to the direct test, which is a likelihood ratio test. However, the estimated residuals of the least squares regression of $y_{t-1}$ on $x_t$ and $x_{t-1}$ are now serially correlated. The estimated serial correlation coefficient of the residual tends to

$$r(1) = \frac{\sigma_v^2 D^{-1}[\rho_x(1) + D^{-1}\sigma_x^2(1) - \alpha - D^{-1}\rho_x(1)\sigma_x(2)]}{1 + \sigma_v^2 D^{-2}\sigma_x^2[1+\rho_x^2(1)] - 2\sigma_v^2 D^{-1}[1+D^{-1}\rho_x(1)\sigma_x^2(1)]},$$

where

$$\sigma_x^2 = \frac{\delta^2\sigma_u^2}{(1-\alpha^2)(1-\gamma^2)} \cdot \frac{1+\alpha\gamma}{1-\alpha\gamma} + \frac{\sigma_v^2}{1-\alpha^2},$$

$$\sigma_x^2(1) = \frac{\delta^2\sigma_u^2(\alpha+\gamma)}{(1-\alpha^2)(1-\gamma^2)(1-\alpha\gamma)} + \frac{\alpha\sigma_v^2}{1-\alpha^2},$$

$$\sigma_x^2(2) = \frac{\delta^2\sigma_u^2(\alpha^2+\gamma^2)}{(1-\alpha^2)(1-\gamma^2)(1-\alpha\gamma)} + \frac{\alpha^2\sigma_v^2}{\alpha^2},$$

$$\rho_x(1) = \frac{\sigma_x^2(1)}{\sigma_x^2}.$$

The usual procedure of re-estimating the model using the r-differenced data will, instead of increasing the power, in general lower the power of the test, thus making comparison of the power of test 3 and test 6 very tricky. The reason that such an adjustment is not necessarily optimal is because under the alternative, the regressors are not independent of the residuals. The problem is similar to the errors-in-variables case analyzed by Grether and Maddala (1973). It might partially explain why DyReyes, Starleaf and Wang (1976) found a reduction in the F-statistic value using the generalized least squares method.

A different problem arises for tests 1 and 2 with the change in the characterization of the driving forces. We note that now $y_t$ no longer possesses a flat spectrum. The lagged cross correlations between the two time series under study are not zero under the alternative hypothesis. The theoretical coherence diagram is no longer a constant. However, it is virtually impossible to compare test 6 with test 2. Test 6 presupposes a known structure of the model, while test 2 takes account of any given lag structure and gives equal weight to all frequency components. Of course, if the structure of the model is known, a direct test has the advantages of essentially weighting each component by its importance and also having better tests available for use with it.

Although a direct comparison between test 2 and test 6 is impossible, a comparison of the asymptotic relative power of test 1 and test 6 may shed some light. We know that under the null, both tests have the same central F distribution. Under the specified alternative, test 6 is a likelihood ratio test, but not test 1 because of the serial dependence of $(\hat{u}_t, \hat{v}_t)$. However, even in such an ideal case for direct test, one fails to establish the optimum property of a regression approach. The criterion we use for evaluating the asymptotic efficiency of a test is the concept of asymptotic relative effiency [Kendall and Stuart (1961, ch. 25)].

Suppose $\tau_1$ and $\tau_2$ are statistics of an hypothesis specified by a parameter $\theta_o$, computed from a sample of size T, their asymptotic efficiency as tests is defined as:

$$(25) \qquad e(\tau_2,\tau_2 | \theta_o) = \lim_{T\to\infty} \frac{\{\frac{\partial}{\partial\theta} E(\tau_1)]_{\theta=\theta_o}\}^2}{Var(\tau_1 | \theta=\theta_o)} \cdot \frac{Var\ (\tau_2 | \theta=\theta_o)}{\{\frac{\partial}{\partial\theta}[E(\tau_2)]_{\theta=\theta_o}\}^2}$$

where $E(\tau_i)$ and Var $(\tau_i)$ are respectively the expected value and variance of $\tau_i$, i = 1, 2. The justification for this criterion is that, under certain regularity conditions, for $\tau_1$ and $\tau_2$ to have the same power against alternative values of $\theta$ which differ from $\theta_o$ by quantities of order $T^{-1/2}$ is in the limit given by $e(\tau_1,\tau_2|\theta_o)$. It can be shown that for test 6

$$(26) \qquad \left\{ \frac{[\frac{\partial}{\partial\delta} E(\hat{\delta})]^2}{Var(\hat{\delta})} \right\}_{\delta=0} = \frac{n\sigma_u^{\ 2}}{(1-\gamma^2)\sigma_v^{\ 2}}$$

While Test 1 is equivalent to test the partial correlation between $x_{t+1}$ and $y_t$ with the effects of $x_t$ removed, thus

$$(27) \qquad \left\{ \frac{\frac{\partial}{\partial\delta} E[r(x_{t+1}y_t | x_t)]^2}{Var[r(x_{t+1}y_t | x_t)]} \right\}_{\delta=0} = \frac{n(1-\alpha^2)\sigma_u^{\ 2}}{(1-\alpha\gamma)^2(1-\gamma^2)\sigma_v^{\ 2}}$$

Therefore the asymptotic relative efficiency for test 1 relative to test 6 is:

$$e = \frac{1-\alpha^2}{(1-\alpha\gamma)^2} \quad . \tag{29}$$

If $\alpha > \frac{2\gamma}{1 + \gamma^2}$, $e < 1$. It follows that even in the ideal case when we know the true structure of the model, it is not necessarily true that a direct test would be more powerful than test 1. Actually, none of the statistics considered so far are asymptotically fully efficient in general. However, a direct test presupposes a structure of the model, but other testing methods do not. Test 3 in the ideal case is a likelihood ratio test, but there is a problem of whether to adjust for serial correlation. Test 1 is basically the same as test 3 except that we choose a correlation approach which allows us to get around this problem in some sense. On the other hand, test 4 prefilters time series data to eliminate serial dependence, but it is deficient because of the low power under the alternative. Test 2 puts the time series data in the frequency domain, thus allowing us to write down the approximate likelihood function neatly. The test statistic with the information generated by the causal coherence and phase diagram would give a measure of the strength of the causality and the extent of the time lags $Y \rightarrow X$ plotted against frequency. They give a more complete picture of the relationship between the time series than just some summary statistics.

## IV. MONEY AND INCOME CAUSALITY DETECTION

The description of the money-income relationship has been a subject of much debate in economic literature. Standing on one side are the quantity theorists who claim that money or its rate of change tends to "lead" income. (Friedman and Schwarz [1963], Friedman [1970], Brunner and Meltzer [1964], etc.) Standing on the other side are the new viewers who provide explicit examples of the possibilities for non-correspondence between causal ordering and temporal ordering of turning points (Brainard and Tobin [1968], Tobin [1970], etc.)

The extreme critics even maintain that the money supply has no relevant place in the determination of money income. They attribute the association between money supply and income to the demand for money which contains income as an important argument. The money stock is somehow called forth to meet the demand. For instance, Davis (1968), Gramley and Chase (1965), Kareken (1967), etc., contend that cyclical fluctuations of monetary growth cannot be attributed to the behavior of the Federal Reserve authorities. These fluctuations are claimed to result primarily from the behavior of commercial banks and the public.

The monetarists, on the other hand, although agreeing with the general relevance of money and income, are far from reaching a consensus with regard to the impact of economic activity on money supply. The strong view (e.g., L. Anderson [1968]) contends that the behavior of the monetary authorities dominates movements in money income. The weak view (e.g., Cagan [1965], Friedman and Schwartz [1963]) does not exclude feedback from income to the money supply, except that the monetary impulses are considered the major factor accounting for variations in money income.

Such controversies can only be resolved by empirical studies. In this section we use the concepts of feedback and causality as they apply to stochastic processes to derive a time series interpretation of the direction of cause and effect. Hopefully it will shed some light on this empirically difficult problem.

We use seasonally adjusted U.S. quarterly money stock and current dollar measures of GNP from 1947 I to 1976 II, for our analysis. Both M1 and M2 were used as alternative measures of money stock variables. All variables were measured as natural logs. Following Sims (1972), we prefiltered each logged variable using the filter $1 - 1.5L + .5626L^2$. Tests for the direction of causality were then performed using methods 1 to 4 discussed in §2 and §3.

Test 3 uses identically the same method as Sims (1972). The income (or money) variable was regressed on current, eight past lags and four future money (or income) variables together with a constant term, trend and seasonal dummies, to test for income-money (or money-income) feedback. If the four future coefficients are significant by an F-test, then we say that feedback is occurring.

Separate autoregressive processes were fitted to prewhiten the money and income series before constructing test 4. Whether the direction of causality goes one way or another depends upon whether the chi-square statistic based on the first four positive or negative cross-correlations between the residuals of money and income is significant or not.

Tests 1 and 2 use a different methodology. To test for causality direction from money to income, we first regress income on its own part and the constant term to remove its past effect. Then we regress money on current and past incomes and the constant term to remove the effect of income variables on money. Test 1, just like test 4, then computes the chi-square statistics from the cross-correlations between the residuals of these two series to test for the feedback from money to income. Test 2, on the other hand, uses the money series, with the effect of income removed, and the income series, with its past effects removed, to compute the causal coherence and uses the normal distribution as its large sample approximation to determine the significance of the feedback. To test for causality from income to money, we similarly remove past effects of money from money and income series.

The results of these tests were reported in Tables 1 and 2. All these tests indicate that the feedback from income to money, if any, was

TABLE 1: MONEY TO GNP FEEDBACK DETECTION

| | M1 | M2 |
|---|---|---|
| $\phi_1(\chi_4^2)$ | 11.097838* | 29.3291** |
| $\phi_2(N(0,1))$ | 1.06383 | 4.6479** |
| $\phi_3(F_{4,86})$ | 7.81875** | 14.1563** |
| $\phi_4(\chi_4{}^2)$ | 8.76262 | 8.39573 |

* Significant at 5% level.

** Significant at 1% level.

TABLE 2: GNP TO MONEY FEEDBACK DETECTION

| | M1 | M2 |
|---|---|---|
| $\phi_1(\chi^2_4)$ | 1.68625 | 3.92849 |
| $\phi_2(N(0,1))$ | -0.294991 | -0.749216 |
| $\phi_3(F_{4,86})$ | 0.011231 | 0.936669 |
| $\phi_4(\chi^2_4)$ | 0.35256 | 4.39254 |

* Significant at 5% level.

** Significant at 1% level.

extremely weak, hence confirming Sims' (1972) earlier finding. However, the test for causality from money to income is less conclusive. Tests 1 and 3 confirm Sims earlier result that causal relation runs from M1 to income, but not for tests 2 and 4. The reason that these results are different might be explained by the observation mentioned in the introduction. That is, if the residual money and income data were not serially independent, tests 1 and 3 might be in favor of the alternative hypothesis, hence generating spurious correlations. On the other hand, tests 2 and 4 take explicit account of the serial dependence of the time series data, thus proper precautions were taken against spurious dependence. However, as demonstrated in §3, the power of test 4 is low, thus will in general favor the null hypothesis of no relations. But test 2 does not have this deficiency. That is why test 2, as well as tests 1 and 3, shows that feedback from M2 to income is occurring, but test 4 is not.

Figures 1 to 10 also provide a heuristic argument in favor of the assertion that causality runs from M2 to income only. These figures of causal coherence, phase diagram, gain and residual spectra of money and income gave a more complete picture about the lead-lag relation between two time series and their strength, which were used to derive our summary statistics. For instance, figures 1 to 5 gave the sample measures of causal coherence, phase, gain, and the residual spectra of GNP and M2 after the effects of past M2 were removed. Figures 6 to 10 gave the similar measures with regard to M2 and GNP after removing the past effects of GNP. A comparison of causal coherence squared (fig. 1 and 6) and the gain from GNP to M2 and M2 to GNP (fig. 3 and 8) indicate that the net effect of M2 to GNP is much stronger than the net effect of GNP to M2. Figure 7 indicates that the causal lead from M2 to GNP is consistent, but figure 2 shows a noisy lead-lag relationship from GNP to M2. Furthermore, figures 9 and 10 indicate that even after removing past effects of income variables, the residual spectra of money and income do not approximate white noise processes. On the other hand, figures 4 and 5 indicate that once the past effects of M2 were removed, the residual M2 spectrum approximates a white noise process, but not the residual GNP spectrum, thus favoring the unidirectional causal relationship between M2 and GNP.

On the basis of tables 1 and 2 and the information provided by the causal coherence, phase, gain and the residual spectra, we may confirm the monetarists' contention that the proper definition of money is M2 and that money or its rate of change "lead" income in some sense.

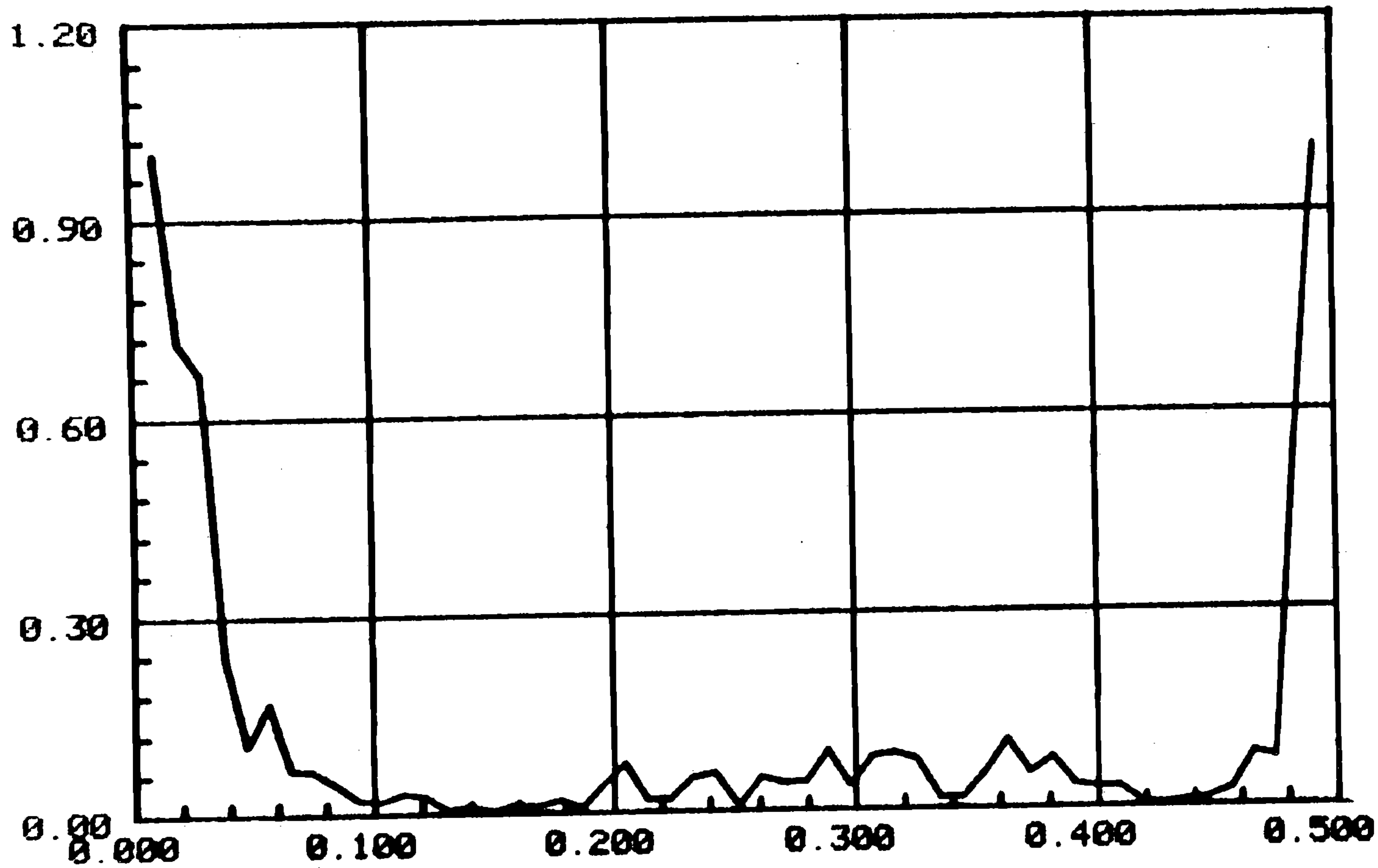


Figure 1: Causal Coherence Squared of GNP to M2

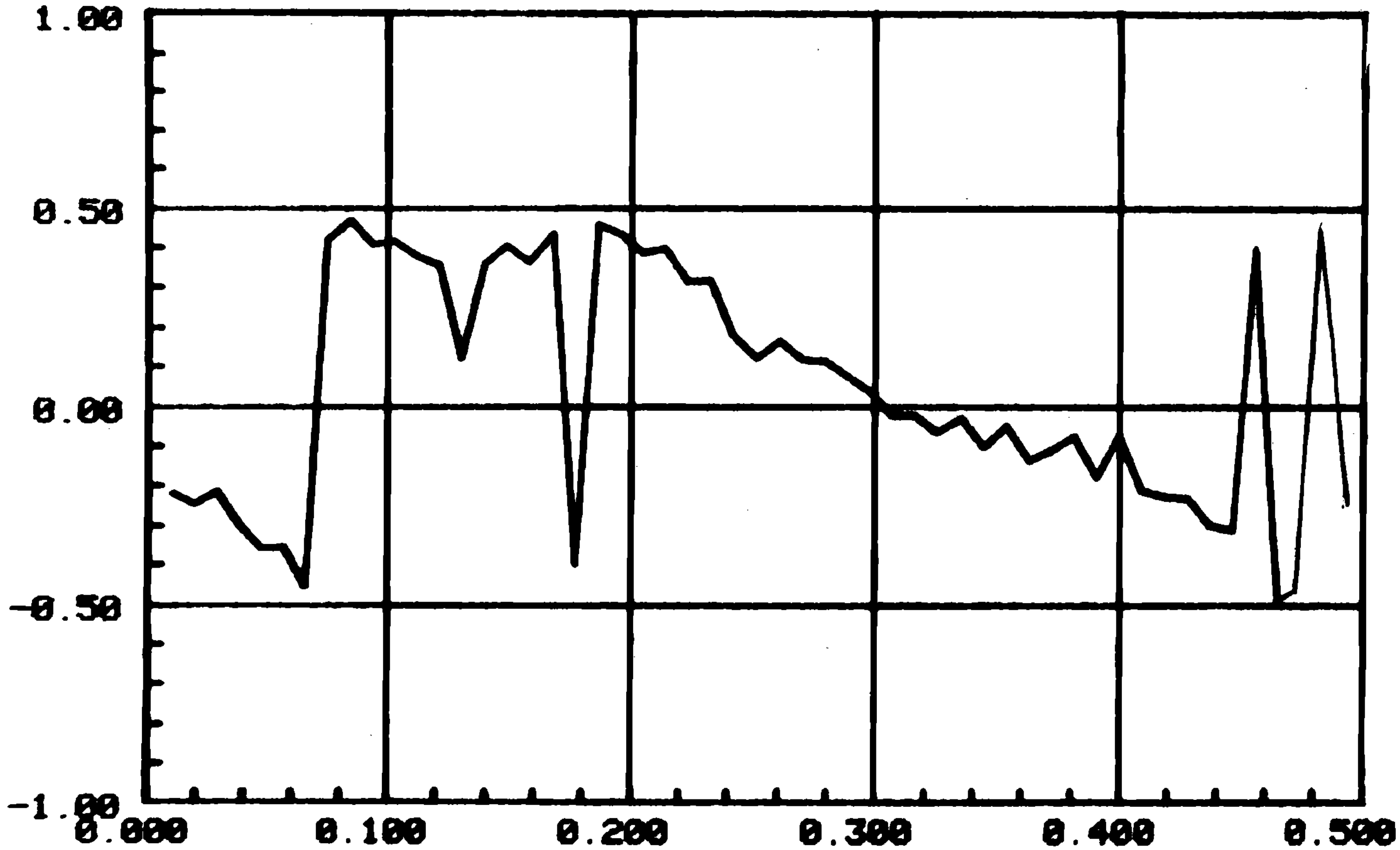


Figure 2: Causal Phase Diagram of GNP to M2

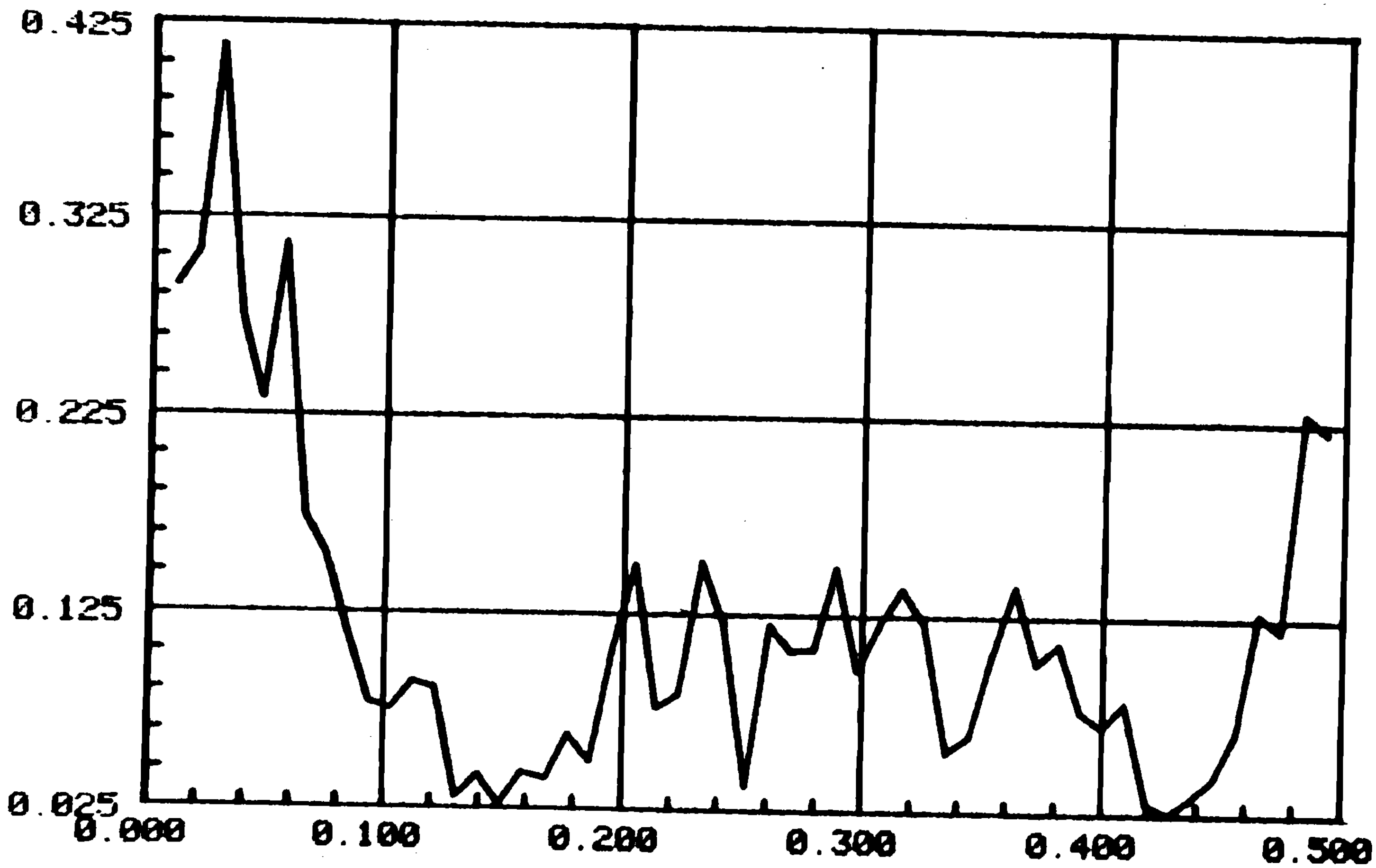


Figure 3: Causal Gain of GNP to M2

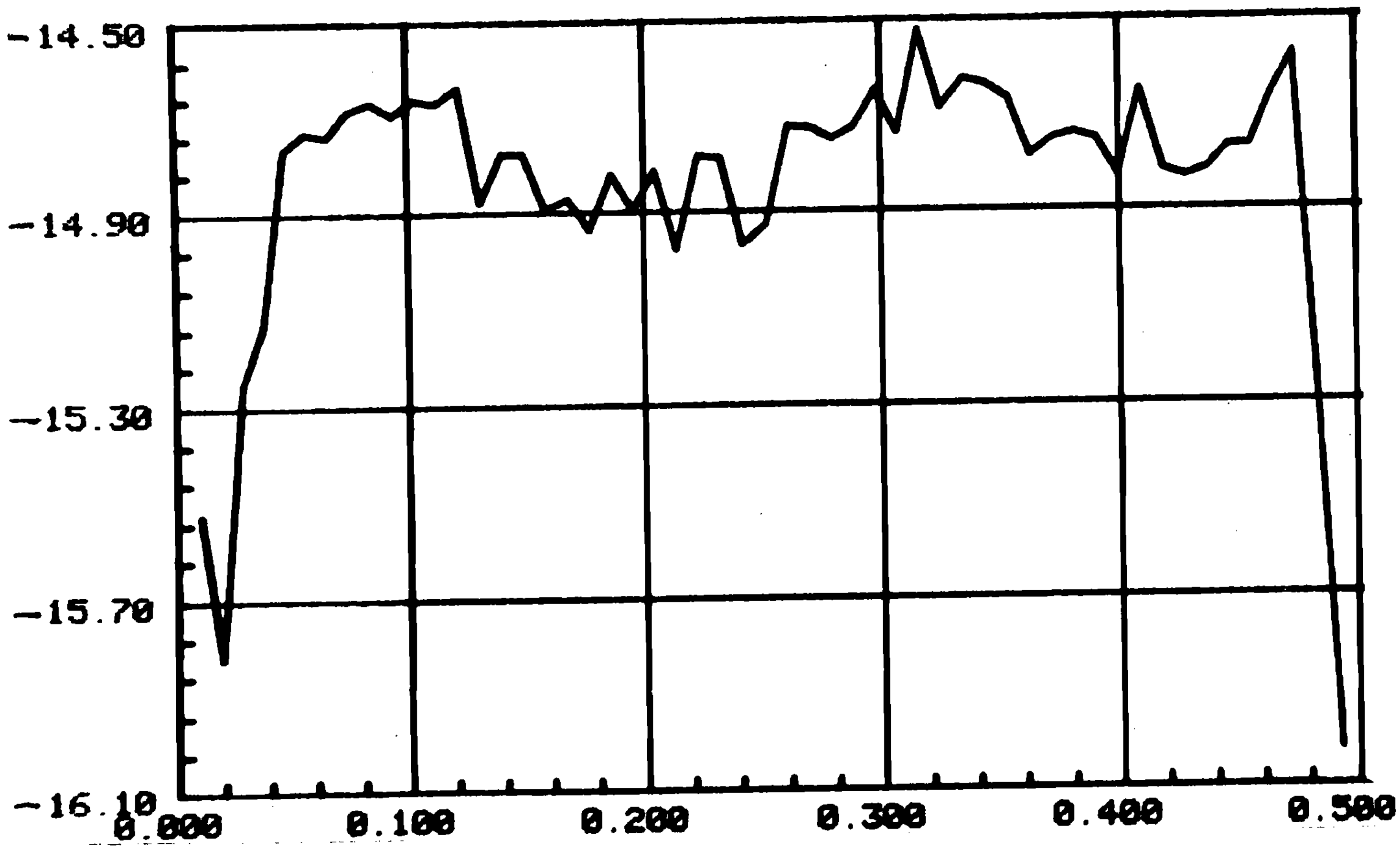


Figure 4: Log-spectrum of M2 after removing the effects of past M2

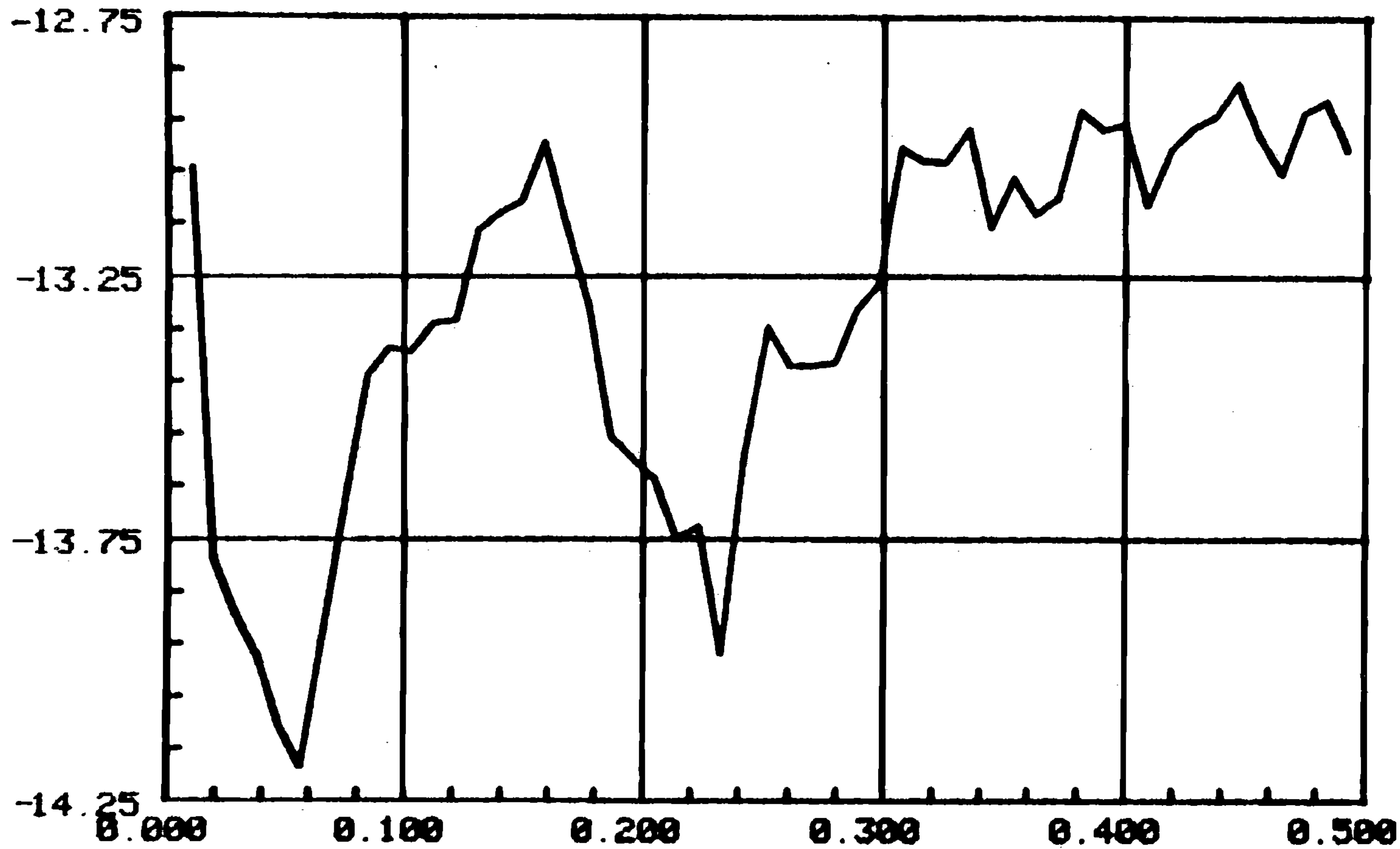


Figure 5: Log-spectrum of GNP after removing the effects of M2

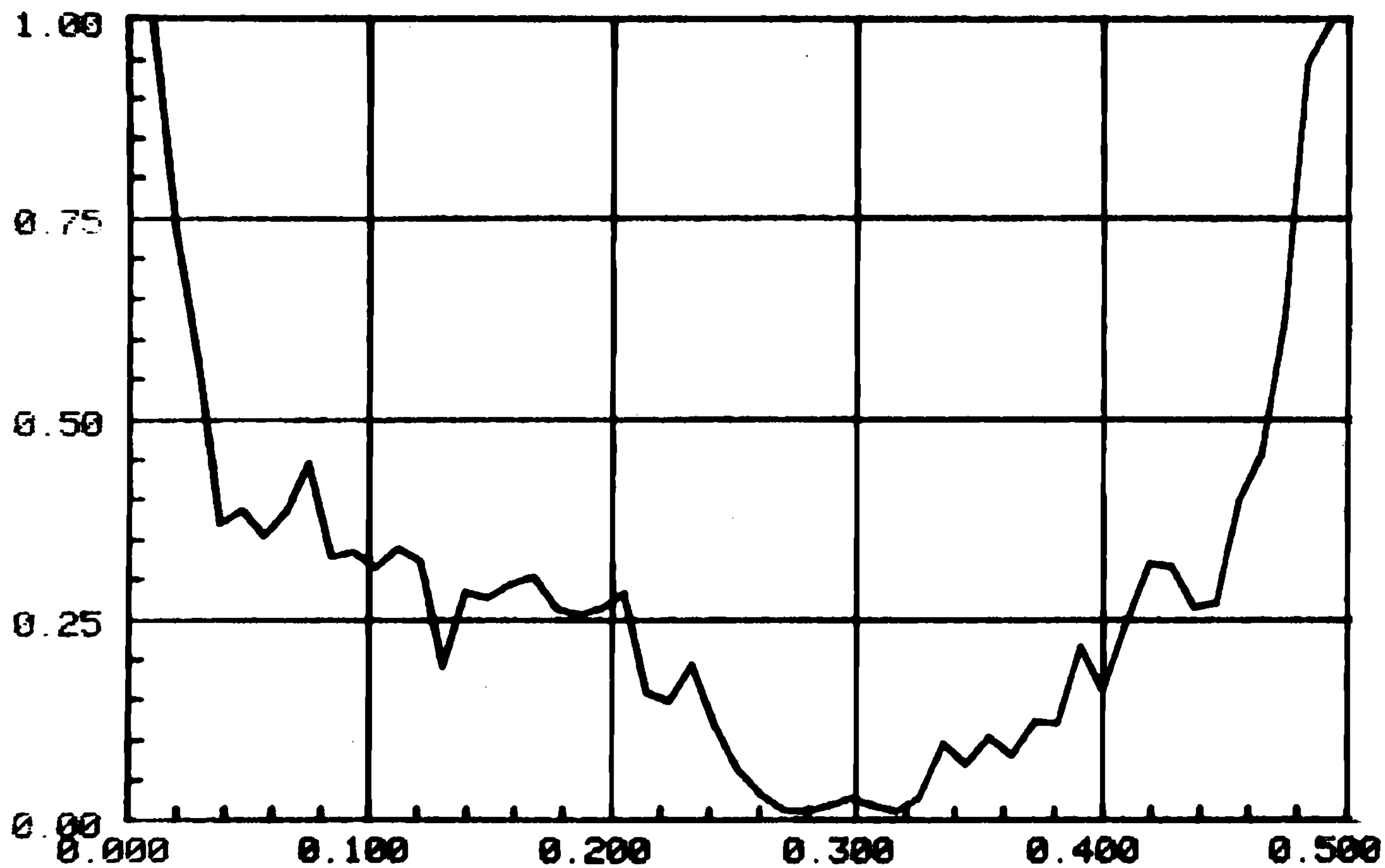


Figure 6: Causal Coherence Squared of M2 to GNP

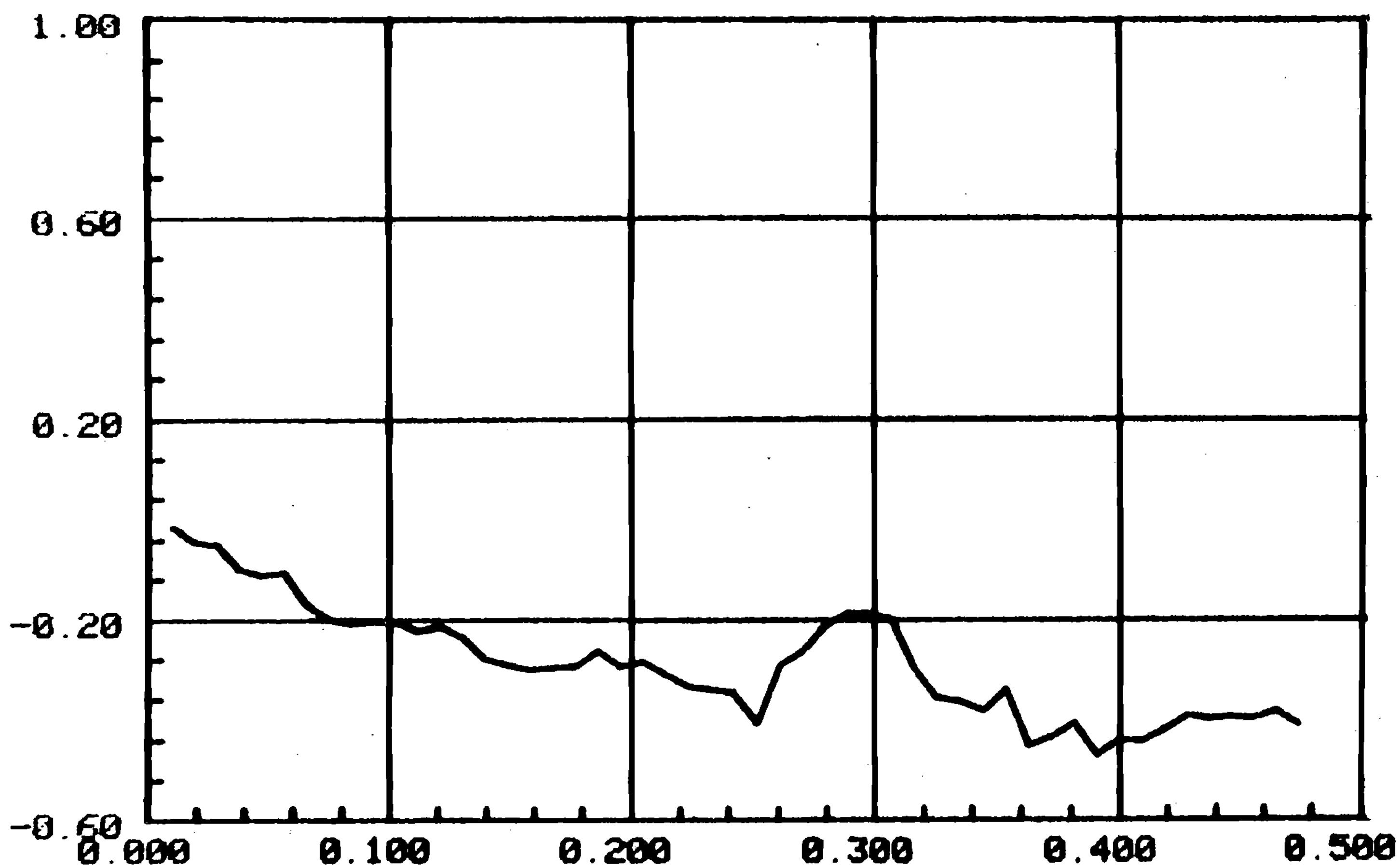


Figure 7: Causal Phase Diagram of M2 to GNP

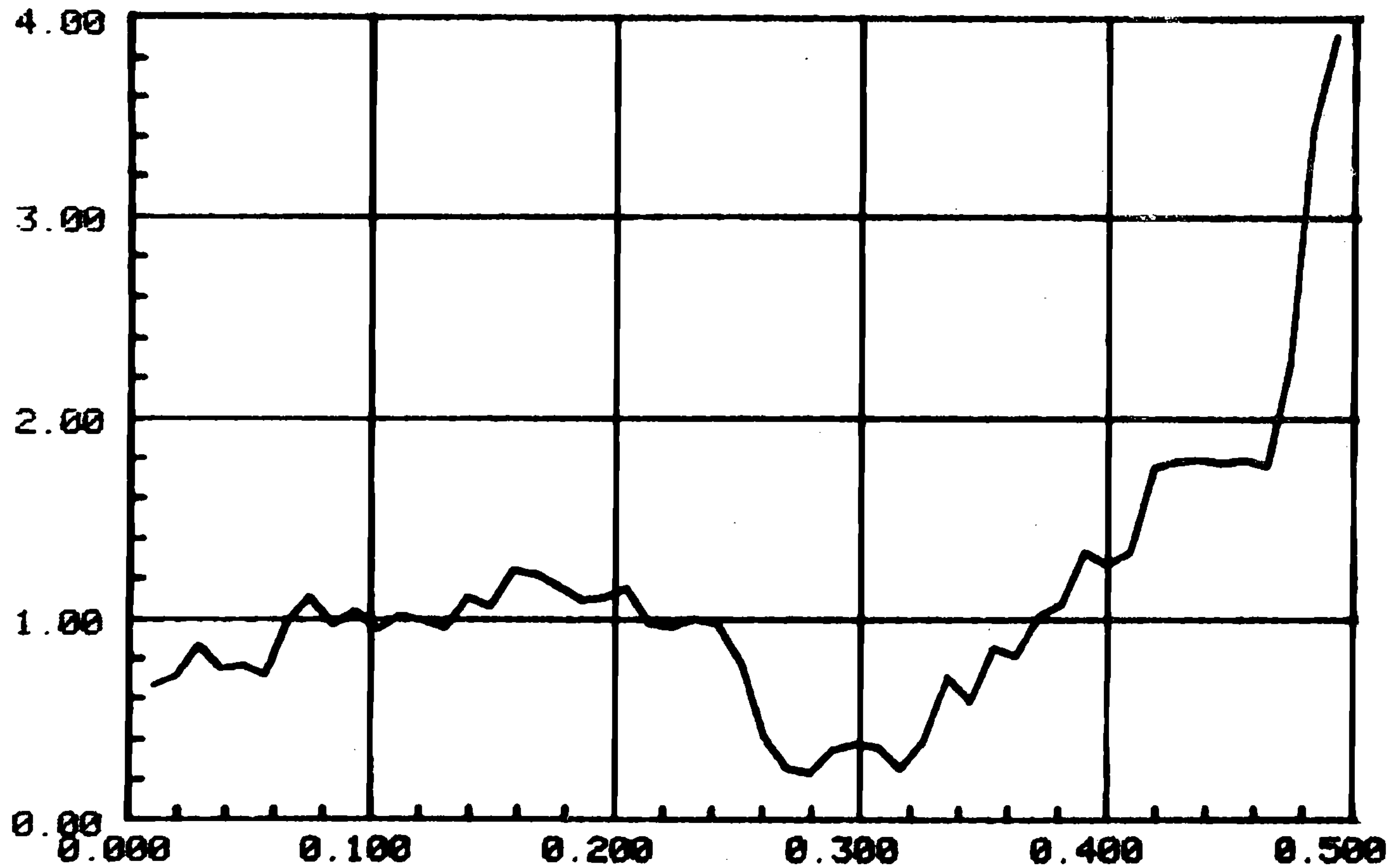


Figure 8: Causal Gain of M2 to GNP

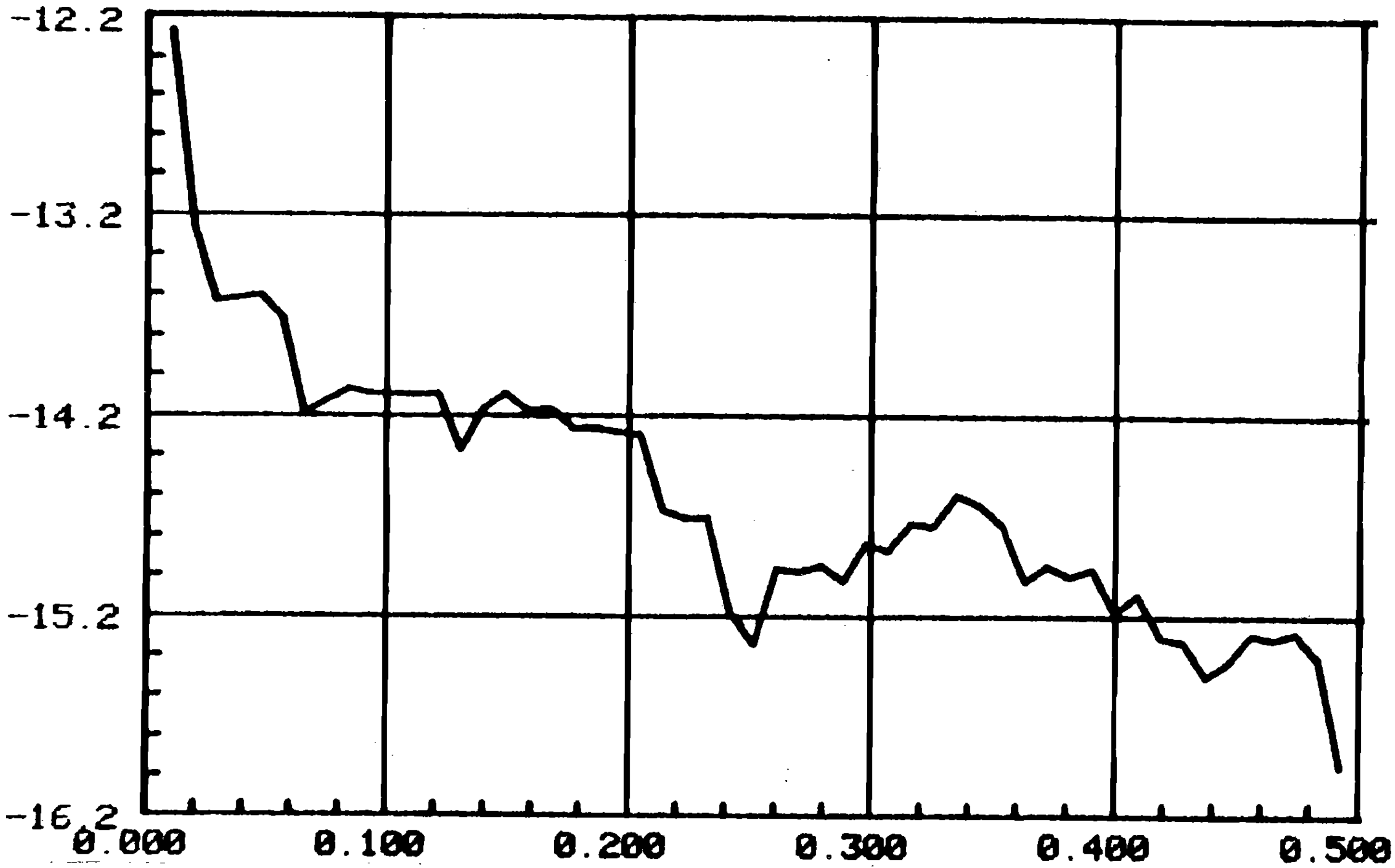


Figure 9: Log-spectrum of M2 after removing the effects of GNP

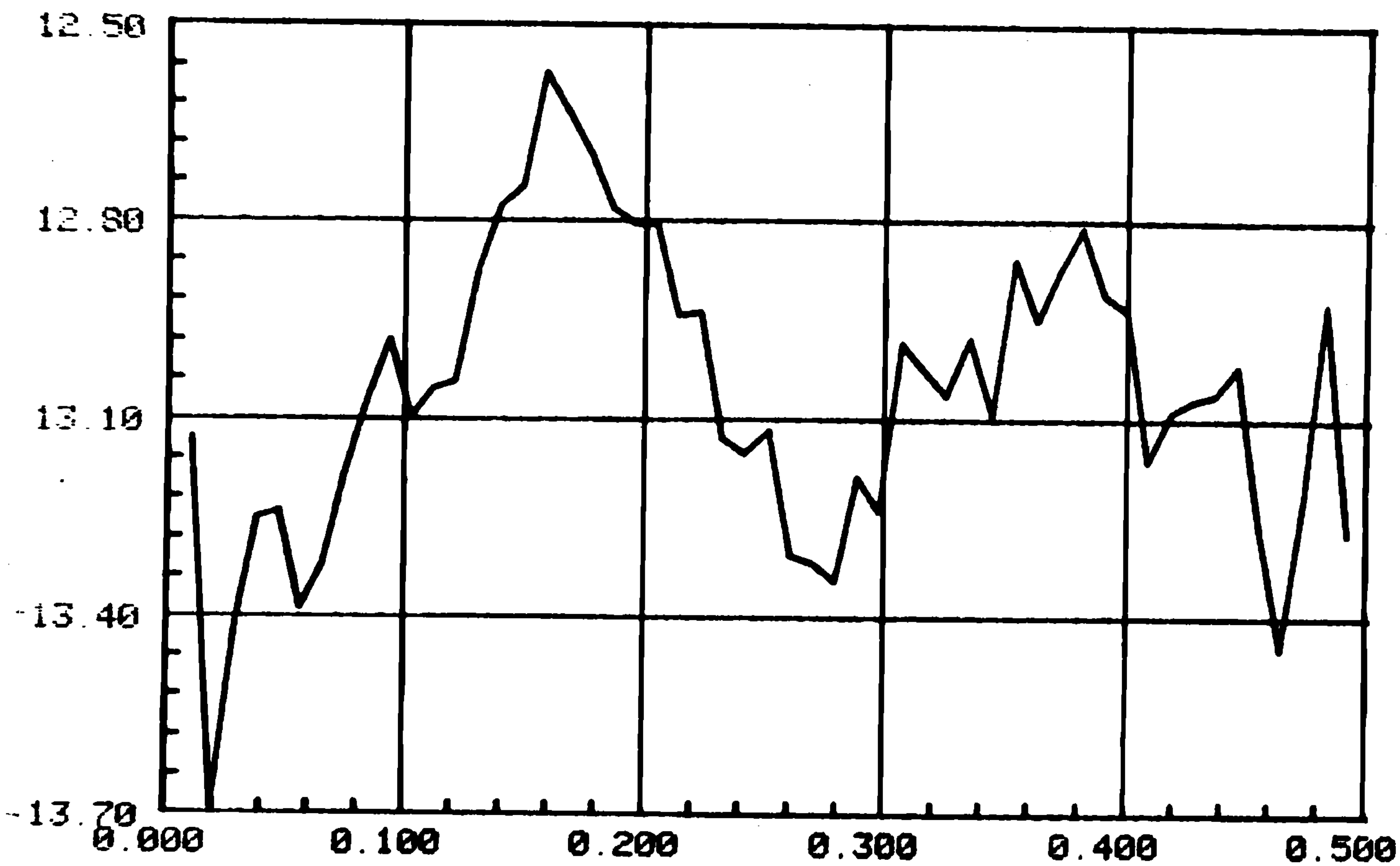


Figure 10: Log-spectrum of GNP after removing effects of past GNP

## V. CONCLUSION

In this paper we summarized the basic theorem characterizing causality events and surveyed existing testing methods. We also presented two alternative testing procedures. One used partial correlation analysis, the other suggested a causal coherence likelihood ratio test. The latter has the advantage of eliminating spurious correlation without the deficiency of losing the power of the test. It is felt that if one suspects that feedback rather than unidirectional causality is occurring in the model, the correlation approach may be an appropriate one to use. Since the regression analysis presupposes the dependence of one or more variables upon others, while feedback is essentially a problem of interdependence, an interest in the joint distribution of a pair of random variables. Furthermore, it provides a unified approach towards the problem of adjustment for serial dependence, which the univariate regression approach does not.

In section 3 we offered an explanation of why various causality detection methods might yield different conclusions. Section 4 applied these methods to the empirical detection of causal direction between money and income. We found that the relationship between M1 and GNP was dubious, but the relationship between M2 and GNP was stable. Both the summary statistics and the spectra diagrams indicated that there was a strong unidirectional influence from M2 to GNP.

Based on the experience of analyzing money and income relationship and the theoretical discussions in section 2 and 3, we feel that if the sample size is small, and the residuals were serially uncorrelated, Sims test is a powerful one to use and is easy to implement. If the sample size is large, then causal spectral analysis would be a powerful exploratory tool. Of course, parametric estimation should be regarded as the ultimate goal in this kind of work. However, the empirical identification of multivariate autoregressive-moving average processes may not be feasible in many occasions. (See Akaike (1974) for some theoretical discussions and Caines and Chan (1975), Wall (1974) for some actual fitting). The main value of causal spectral analysis is in its flexibility in examining the sources of variations in the data. It can be used to suggest possible models.

# FOOTNOTES

1. The result of their method III only.

2. These definitions correspond to the weak feedback-free (WFF) process given by Caines (1976). Cains also gave a corresponding definition on causal process pairs without instantaneous causality (Granger [1969], Pierce and Haugh [1977]), which he refers to as strong feedback-free process (SFF). All the testing procedures to be discussed below can be easily modified to test for SFF. However, since it is difficult to say who is causing whom when instantaneous causality was observed and the strong belief by this author that should there be a causal relation between two time series variables, the inertia will make it last into future, we choose the WFF definition.

3. If we wish to test for strong feedback-free, we regress $y_t$ on past $x_t$'s only and sum $\hat{r}^2_{uv}(k)$ from k=0 to k=M in (7) and check for the critical region of chi-square statistic with (M+1) degrees of freedom.

4. The (T-p) here is not an adjustment of degrees of freedom, but because we have only (T-p) estimated residuals.

5. For a precise condition, see Brillinger (1975) or Wahba (1968).

6. I feel compelled to compare these various tests at its ideal setup with the given prior information. Otherwise, there are too many alternatives to choose from.

# References


Akaike, H. (1974) "A New Look at the Statistical Model Identification." IEEE on Automatic Control, AC-19, 716-723.

Anderson, L. (1968) "Monetary and Fiscal Actions: A Test of Their Relative Importance in Economic Stabilization." Federal Reserve Bank of St. Louis Monthly Review 50 (November 1968), 11-23.

Anderson, T.W. (1958) An Introduction to Multivariate Analysis. Wiley, New York.

__________. (1971) The Statistical Analysis of Time Series. Wiley, New York.

Box, G.E.P. and D.A. Pierce. (1970) "Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models." Journal of the American Statistical Association 65, 1509-1526.

Brainard, W. and J. Tobin. (1968) "Pitfalls of Financial Model Building." American Economic Review Proceedings 58, 99-122.

Brillinger, D.R. (1975) Time Series: Data Analysis and Theory. Holt, Rinehart and Winston, San Francisco.

Brunner, K. and A.H. Meltzer. (1966) "A Credit Market Theory of the Money Supply and an Explanation of Two Puzzles in U.S. Monetary Policy." Essays in Honor of Macro Fanno. Padova, Italy.

Cagan, P. (1965) Determinants and Effects of Changes in the Stock of Money. National Bureau of Economic Research, Inc., New York.

Caines, P.E. (1976) "Weak and Strong Feedback Free Processes." IEEE Trans. Automatic Control, AC-21, 737-739.

__________ and C.W. Chan. (1975) "Feedback between Stationary Stochastic Process." IEEE Trans. Automatic Control, AC-20, 498-508.

Chitturi, R.V. (1974) "Distribution of Residual Autocorrelations in Multiple Autoregressive Schemes." Journal of the American Statistical Association 69, 928-934.

Davis, R.G. (1968) "The Role of the Money Supply in Business Cycles." Monthly Review, FRB of New York (April 1968).

DyReyes, F.R. Jr., D. Starleaf and G.H. Wang. (1976) "Tests of the Direction of Causation Between Money and Income in Developed Countries." Paper presented at the North American Meetings of the Econometric Society, Atlantic City, New Jersey.

Feige, E. and D.K. Pearce. (1974) "The Causality Relationship Between Money and Income: A Time Series Approach." Paper presented at the Midwest Economic Association Meetings, Chicago, April 1974.

Friedman, M. (1970) "A Theoretical Framework for Monetary Analysis," in Milton Friedman's Monetary Framework, ed. by R.J. Gordon, 1-62. Chicago: Chicago University Press.

___________ and A. Schwartz. (1963) "Money and Business Cycles." Review of Economics and Statistics, Supplement, 45, 32-64.

Gramley, L. and S. Chase. (1968) "Time Deposits in Monetary Analysis." Federal Reserve Bulletin (October 1965).

Granger, C.W.J. (1969) "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods." Econometrica 37, 424-438.

___________ and P. Newbold. (1974) "Spurious Regressions in Econometrics." Journal of Econometrics 2, 111-120.

Grether, D.M. and G.S. Maddala. (1973) "Errors in Variables and Serially Correlated Disturbances in Distributed Lag Models." Econometrica 40, 255-262.

Hannan, E.J. (1970) Multiple Time Series. Wiley, New York.

Haugh, L.D. (1976) "Checking the Independence of Two Covariance-Stationary Time Series: A Univariate Residual Cross-Correlation Approach." Journal of the American Statistical Association 71, 378-385.

Kareken, H. (1967) "Commercial Banks and the Supply of Money: A Market Determined Demand Deposit Rate." Federal Reserve Bulletin (October 1967).

Kendall, M.G. and A. Stuart. (1961) The Advanced Theory of Statistics, Volume II. Hafner Publishing Co., New York.

Pierce, D.A. (1977) "Relationships -- and the lack thereof -- between economic time series, with special reference to money, reserves, and interest rates." Journal of the American Statistical Association, forthcoming.

Pierce, D.A. and L.D. Haugh. (1977) "Causality in Temporal Systems: Characterization and a Survey." Journal of Econometrics, forthcoming.

Quenoulli, M.H. (1948) "Approximate Tests of Correlation in Time Series." Journal of Royal Statistical Society, Supplement, 11, 68-84.

Rao, C.R. (1962) "Efficient Estimates and Optimum Inference Procedures in Large Samples." Journal of Royal Statistical Society,(B) 24, 46-72.

Sims, C.A. (1972) "Money, Income, and Causality." American Economic Review LXII, 540-552.

Tobin, J. (1970) "Money and Income: Post Hoc Ergo Propter Hoc?" Quarterly Journal of Economics 84, 301-317.

Wahba, G. (1968) "On the Distribution of Some Statistics Useful in the Analysis of Jointly Stationary Time Series." The Annals of Mathematical Statistics 39, 1849-1862.

Wall, K.D. (1974) "An Application of Simultaneous Estimation to the Determination of Causality Between Money and Income." Discussion Paper No. 8, Programme for Research into Econometric Methods, University of London.

Williams, D., C.A.E. Goodhart, and D.H. Gowland. (1976) "Money, Income, and Causality: The U.K. Experience." American Economic Review, 417-423.

Wu, D.M. (1973) "Alternative Tests of Independence Between Stochastic Regressors and Disturbances." Econometrica 41, 733-750.