# EXOGENEITY TESTS AND MULTIVARIATE TIME SERIES: PART 1

by

Christopher A. Sims

Discussion Paper No. 75 - 54, June 1975

Center for Economic Research
Department of Economics
University of Minnesota
Minneapolis, Minnesota 55455

EXOGENEITY TESTS AND MULTIVARIATE TIME SERIES: Part 1

This paper takes up a number of issues which arise when tests like those used in "Money, Income, and Causality" [9] are applied in the context of either a system of simultaneous behavioral equations or a "non-structural" vector autoregressive-moving-average (ARMA) model. Some of these same issues have been taken up by Zellner and Palme [10] and by David A. Pierce and Larry P. Haugh [7].

Testing for exogeneity in behavioral simultaneous equation models or in subsets of equations from such models is possible by methods directly analogous to those of [9].

Using such methods, it is not necessary to embed a behavioral model in a complete vector ARMA model in order to formulate an exogeneity test, and the paper points out that there can be advantages in doing exogeneity tests without explicit modeling of the process determining the hypothetically exogenous variables.

The paper goes on to show how a simultaneous equation model can be embedded in an autoregressive-moving-average vector stochastic process model, generalizing the discussion of this topic by Zellner and Palme. This analysis makes clear a sense in which an exogeneity test is a test for the presence of simultaneity bias.

A number of researchers have applied exogeneity tests of a form which does require estimation of the process determining the exogenous variables. Legitimate methods exist for doing this, and such methods have advantages in some applications. However, some apparently natural procedures for executing such tests are in fact unjustifiable, though they have been used in practice.

Part II of this paper (still to be written as of May, 1975) takes up questions related to the behavioral or "economic" interpretation of the results of exogeneity tests.

## 1. Exogeneity in Behavioral Models

Assume that a system of behavioral equations can be written

1) $$A(L)Y + B(L)X = C(L)U ,$$

where A, B, and C are finite-order matrix-valued polynomials in non-negative powers of the lag operator $L$. The error terms $U$ are assumed to have a mean of zero and to be serially uncorrelated. Note that if the error terms in the original behavioral specification have an autoregressive component in their serial correlation strucutre, then $A$ and $B$ may not be what we usually think of as behavioral parameters. For example, suppose we start with a system $\bar{A}(L)Y + \bar{B}(L)X = V$ , $F(L)V = U$. Then to get a system of the form (1) we have to "multiply" through by F(L), so that $A$ and $B$ in (1) correspond to $F(L)\bar{A}(L)$ and $F(L)\bar{B}(L)$.

Suppose that under the assumption that $X$ is strictly exogenous, the system (1) is identified, so that we can undertake inference about its parameters. The assumption of strict exogeneity, as it is used in the Gauss-Markov theorem and in estimation methods based on it, takes the form

2) $$E[U(t_o) | X(t_1),\ldots X(t_k)] = 0 ,$$
for any $\{t_o, t_1, \ldots t_k\}$.[1/]

A test is available for the hypothesis that X is exogenous: Expand the model (1) to the form

3) $$A(L)Y + B(L)X + G(L)X = C(L)U \quad ,$$

where G involves only negative powers of L, and test the null hypothesis that all the coefficients in G are zero. If (1) is the true model with X in fact exogenous, then the coefficients in G are zero. The polynomial G could include positive powers of L higher than the degree of B, but unless we are in the rare situation of having a priori theoretical constraints on the degree of B, an econometric model will generally have been chosen by trial and error to guarantee that no higher powers of L than appear in B will possess statistically significant explanatory power. We could also include in G positive powers of L less than or equal to the degree of B, but coefficients on these terms are

---

[1] This is actually a slightly stronger assumption than is required for the Gauss-Markov theorem. That theorem would require only that (2) hold for any collection $\{t_o, t_1, \ldots t_k\}$ all of whose elements are contained within the sample period. In time series, if the assumption is not to be dependent on sample length or on absolute location in time of the sample, (2) must hold for arbitrary $\{t_o, t_1, \ldots t_k\}$.

likely not to be identified. In an overidentified model, coefficients on some such terms will be identifiable, and then G could certainly include such terms.

Testing for exogeneity this way requires that (1) be identified and remain so when expanded to (3) with an appropriate parameterization of G. It is not required that the system (1) be complete -- the number of equations could be less than the dimensionality of Y. Furthermore, it is not required, even if A(0) is of full rank, that A(L) have inverse expressible as a polynomial in positive powers of L with convergent coefficients. That is, it is not required that A be "stable".[2/]

---

[2]That A be stable when A(0) is square may seem an innocuous assumption. However, in the first place there are applications where gently nonstationary behavior for Y (e.g., growth at less than 10% per year) may be reasonable, and a system which generates that behavior endogenously might involve unstable A. In the second place, Y and X might both be stationary, even though A were "unstable". For example (this example is inspired by unpublished work by Peter K. Clark of Minnesota and John Kennan of Brown) we might have a model in which

$$V(t)-W(t) = (1/r)[(V(t-1)-W(t-1)) - P(t-1)] + U(t) ,$$

where V(t) is the share value of a conglomerate firm, W(t) is the share value of a subsidiary corporation of the conglomerate, r is the discount factor, P(t) is earnings per share of the conglomerate, net of earnings of the subsidiary, and U(t) is a residual. Under certain assumptions, it can be argued that the expected value as of time t of the left-hand side of this equation is the right-hand-side with the error term suppressed. It then follows that U is serially uncorrelated and uncorrelated with the variables appearing on the right-hand side of the equation, so the equation can be estimated consistently by ordinary least squares. Under the further assumption that the fortunes of the subsidiary and of the remainder of the conglomerate are independent, W is exogenous with respect to U and

If $A(L)$ is square and we know that $\det(A(L))$ has all its roots strictly outside the unit circle of the complex plane, then an infinite polynomial in positive powers of $L$ with absolutely summable coefficients, called $A^{-1}(L)$, exists such that $A^{-1}(L)A(L) = I$. In this case a procedure for testing exogeneity of X exists even when (1) is not identified. Multiplying (1) through by $A^{-1}$, we obtain

4) $$Y = -A^{-1}(L)B(L)X + A^{-1}(L)C(L)U \ .$$

The system (4) is a set of distributed lag relations with (on the null hypothesis) current and lagged values of strictly exogenous variables on the right-hand side and serially correlated errors In the rare situation where we have a priori knowledge of the orders of A, B, and C we can obtain a finite-dimensional parameterization of (4) directly and proceed with estimation. More generally, it will be as easy to choose a priori finite lengths for the lag distributions in an approximation to (4), such that

---

that hypothesis can be tested as suggested above. However if we add to the system an equation explaining P as a function of exogenous variables and past values of itself, the resulting two-equation system has an "unstable" A. This does not imply that V-W or P must be non-stationary. It does imply that (V(t)-W(t), P(t)) does not lie in the space spanned by current and past values of the residuals in the two-equation system and of the exogenous variables. Intuitively, the system is not one which "determines" V-W and P from the history of exogenous variables and error terms, though it is nonetheless a valid pair of behavioral equations.

the omitted tails of the lag distributions contribute explanatory power of trivial variance. In either case, once estimates of (4) are obtained, future values of X can be added to the right-hand-side of (4) and the null hypothesis that their coefficients are zero tested by standard methods. This null hypothesis is implied by the hypothesis that X is exogenous, and testing it is therefore a test of exogeneity.[3]

## 2. Exogeneity in Vector ARMA Models

If the exogenous variables X in (1) are stationary, then it is assured that X is representable to an arbitrarily close [4] approximation by a vector ARMA

---

[3] In my earlier paper [9] I emphasized the connection between the null hypothesis of strict exogeneity and Granger's [3] definition of a one-way causal ordering. This connection is real, but it is potentially misleading as well, because of the variety of ways we use the word "cause". We might estimate a single equation distributed lag investment accelerator model without believing that "output causes investment" in some legitimate sense of the work "cause"; but if we are using estimation methods predicated on strict exogeneity, we should nonetheless test for exogeneity of output in the estimated equation. In fact, Hannan [4] had pointed out years ago that exogeneity tests were possible in principle, and Phillips [6] discussed, even more years ago, the relevance of exogeneity (and generalizations thereof!) to identification of systems of time series relations.

[4] This approximation can be made arbitrarily close in the sense that the spectral density of X implied by (5) can be made arbitrarily close to the true spectral density in the $L_1$ metric by picking the degree of H and J high enough. This implies that the joint covariance matrix of $X(t+s_1)$, $X(t+s_2),\ldots,X(t+s_k)$ for fixed $s_1, s_2,\ldots,s_k$ implied by (5) can be made arbitrarily close to the true covariance matrix.

structure of the form

$$5) \qquad H(L)X = J(L)W ,$$

where W is a vector of mutually uncorrelated white noises, His "stable", and H and J are finite-order polynomials in nonnegative powers of L with square matrix coefficients. If we append (5) to (1), if A is square, if Y and X are jointly stationary, and if X is exogenous in (1), then (1) and (4) together form an ARMA representation for the vector process (Y, X). Exogeneity can be tested in this system by testing the null hypothesis that W and U are uncorrelated, and it should be clear that adding G(L) to (1) to get (3) and testing the hypothesis G = 0 is a test of the hypothesis that U and W are uncorrelated.[5/]

The class of tests for exogeneity proposed in Section 1 above make no use of (5). This may be an advantage or a disadvantage. If we begin with the system (1) or (4), have some a priori knowledge (possibly imprecise) of the form of one of those systems, and have no a priori knowledge of the form of (5), then it is certainly an advantage to avoid using (5) in forming a test of the exogeneity of X in (1). On the other hand we might have overidentifying

[5]This condition, that if Y and X are jointly stationary and X is exogenous in a behavioral system (1), then we can form a vector ARMA representation for (Y,X) by appending (5) to (1) with U and V mutually uncorrelated, has been noted previously by Zellner and Palme ([10], p. 19-20). As Zellner and Palme state the condition, however, the reader might gain the erroneous impression that if an estimated ARMA representation fails to meet the condition, then X cannot be exogenous in a behavioral model explaining Y. In fact, since the ARMA representation is not unique, the condition is in a sense only necessary. This point is taken up again in 6.

restrictions on both (1) and (5).[6] In this case there would be an advantage in testing the system formed by (1) and (5) against a more general, but still identified, ARMA model in which new terms in W appear in (1), new terms in U appear in (5), and/or new terms in Y appear in (5). Such a test does test exogeneity of X in (1), but of course it also tests all the other overidentifying restrictions at the same time. In heavily overidentified systems, or where identification has been achieved by the use of assumptions more dubious than exogeneity of X in (1), there is not much point in labeling a test like this an "exogeneity test".

Finally it is possible that (5) is identified and (1) is not.[7] To generate a test for exogeneity of X in (1) when (1) is not identified we need to know that X and Y are jointly stationary and that A can be chosen square and with all the roots of $\det(A(L)) = 0$ outside the unit circle. In that case Y can be expressed as a linear combination of current and past values of X and U, so on the null hypothesis that X is exogenous in (1) values of Y(t-s) for s greater than zero are uncorrelated with W(t). This hypothesis can be tested by standard methods.

Even if neither (1) nor (5) is separately identified,

---

[6] For example, (5) might be a simultaneous equation model of a set of macro-economic aggregates and (1) might be a model of the construction industry in which the macro-economic aggregates are taken as exogenous.

[7] E.g., in our construction-industry-model example we might have a complete model for the aggregate variables X already available, and merely be contemplating construction of the model(1) for the construction industry. As a preliminary step we might wish to ask whether it is possible that X could be exogenous in a model like (1).

it is possible to test the null hypothesis that X is exogenous in a system of the form (1). One way to do this which does not depend on X being stationary or conveniently modeled as a vector ARMA has already been described: Testing for non-zero coefficients on future X in (4). If X is stationary and has some exact representation of the form (5), then H can be chosen "stable" and J can be chosen with all the roots of det J(L) = 0 on or outside the unit circle (stable or borderline unstable).[8/] Also, if X is stationary, any approximate representation of X in the form (5) can be replaced by one implying exactly the same covariance properties but with H stable and the roots of det J(L) = 0 on or outside the unit circle. These restrictions do not, however, suffice to identify H and J. Hannan [5] gives conditions from which a normalization rule which will identify H and J can be derived.[9/]

---

[8] If (5) represents X , then X is a rational process, and a rational moving average representation can be chosen for it in which the disturbance vector is "fundamental" for X in the sense of Rozanov [8]. For a rational process, W being fundamental for X in (5) is equivalent to the conditions on H and J given in the text.

[9] However, it is in my view a fundamental disadvantage of ARMA representations in situations where the orders of H and J are not known a priori that a conservative "overfitting" procedure (choosing the orders of H and J to be certainly high enough and probably higher than necessary) undermines identification. Since purely autoregressive representations (ARMA representations with J(L)=I) do not suffer from this defect, can equally well approximate an aribtrary stationary process, and are generally easier to estimate than ARMA processes, it will be sensible in many applications to achieve identification by taking the order of J to be zero, the order of H to be conservatively high, and H(0) to be upper triangular.

With (5) identified by some normalization rule, testing of the exogeneity hypothesis proceeds exactly as if (5) were identified as a behavioral system. Lagged values of $Y$ are uncorrelated with $W$ on the null hypothesis, so the null hypothesis can be tested as a linear restriction.

It might seem that instead of first estimating (5), then testing for correlation of $W$ with lagged $Y$'s, one could instead arbitrarily normalize the joint system formed by (1) and (5), testing exogeneity as a restriction on the joint ARMA representation. That is, one might estimate a system of the form

$$\begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} \begin{bmatrix} Y \\ X \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} U_1 \\ U_2 \end{bmatrix} , \tag{6}$$

with $U_1$ and $U_2$ mutually uncorrelated white noises, $A_i$, $B_i$, $C_{ij}$, finite-order polynomials in nonnegative powers of $L$ for all i, j, and test the exogeneity of $X$ as a restriction on (6).

But what is the restriction on (6) implied by exogeneity of $X$ ? Since (1) and (5) together, with $X$ exogenous, form a system like (6) with $A_2=0$, $C_{21}=0$, and $C_{12}=0$, the natural procedure might be to estimate (6) and test these three sets of zero-restrictions. This natural procedure is not a good one.

The system (6) is not uniquely determined representation of (Y,X). The fact that $A_2=0$, $C_{21}=0$, and $C_{12}=0$ in one valid representation of (Y,X) in the form (6) does not mean that in choosing some particular normalization rule to identify a vector ARMA scheme like (6) we will necessarily end up estimating the particular ARMA scheme in which the restrictions hold. It can happen, for example, that even though there is a vector ARMA representation satisfying the restrictions, the vector ARMA representation involving the fewest lags satisfies none of the restrictions.

Consider the following simple system in which Y and X are both scalars:

$$\begin{aligned} &7) \quad Y_t - .3Y_{t-1} &&= U_t - .3U_{t-1} + .4V_{t-1} \\ &\qquad Y_t + X_t - .2X_{t-1} &&= U_t + V_t \quad . \end{aligned}$$

The system involves no lags greater than first order, and clearly neither equation as written is an equation in which X can be taken as exogenous. However (7) is equivalent to

$$\begin{aligned} &8) \quad Y_t - Y_{t-1} + .21Y_{t-2} - .4X_{t-1} + .2X_{t-2} - .024X_{t-3} = U_t - U_{t-1} + .21U_{t-2} \\ &\qquad\qquad X_t - .5X_{t-1} + .06X_{t-2} = V_t - .7V_{t-1} \quad , \end{aligned}$$

in which all the restrictions are met so that X is exogenous in the first equation. If we began by estimating (7), we could determine rightly that lags no higher than first order in Y, X, U, and V are required to represent the system. If, however, we then tested the restrictions implied by exogeneity of X in the first equation while maintaining the hypothesis that lags run no higher than first order, we could reject the null hypothesis even though a relation of Y to current and past X with X exogenous does exist.

The problem is that to choose orders of H and J appropriately we would have to begin with the constrained ARMA system, which involves estimating (5) in any case. Having instead estimated a good-fitting unconstrained low-order vector ARMA representation for (Y,X), an appropriate test for the null hypothesis that a relation between Y and X with X exogenous exists must test the hypothesis that some transformation of the estimated representation would satisfy the exogeneity conditions. This hypothesis is complex and nonlinear and will not be specified in detail here, because there are more convenient test procedures.

## 3. Details and Pitfalls of Exogeneity Testing

One normalization rule which will identify (5) sets $J(0)$ to be the identity and $H(0)$ to be upper triangular with ones down the main diagonal. Maximum likelihood estimation, on the assumption that $J^{-1}$ exists, amounts to minimizing the product of the sum of squared residuals in the residuals in the equations of the system

$$J^{-1}(L)H(L)\ X = W \quad . \tag{9}$$

In the case where J has been chosen to be the identity, each parameter of (9) appears in only one equation and ordinary least squares (OLS) can be applied equation by equation. To test the hypothesis that X is exogenous in (1), under the maintained hypothesis that A in

(1) has a convergent one-sided inverse (A is stable), we expand (9) to become

10) $$J^{-1}(L)H(L)X + K(L)Y = W \quad ,$$

where K involves only strictly positive powers of L (K(0)=0). Under the null hypothesis K must be zero.[9/] In the purely autoregressive parameterization (J(L)=I) this hypothesis is particularly easy to test, since in that case, if separate coefficients are allowed for lagged Y's in the separate equations of (10), the null hypothesis K=0 becomes a set of linear hypotheses in a set of linear equations estimable by OLS.

The test suggested in section 1, adding future values of X to the right-hand-side of (4) and testing the null hypothesis that they have zero coefficients, tests the same null hypothesis as the one that K=0 in (10), if both (1) and (5) hold. The difference between the two tests is that the test based on (10) requires that (5) hold, but not necessarily that Y be stationary or that we know how to specify (1). The test based on (4) requires that

---

[9] The requirement that K(0) be held to zero, i.e. that no contemporaneous Y's be added to the equation in (10), stems from the fact that (1), with stable A, implies that Y is a linear combination of current and past X and current and past U. Since W in (5) is uncorrelated, on the null hypothesis, with U, W is uncorrelated with past Y. But current Y contains a current X component, hence a current W component, and hence is in general correlated with current W on the null hypothesis.

we know how to specify (1), but not necessarily that we know $X$ is stationary or that we know how to specify (5).

It has been suggested in several unpublished papers (e.g. Feige and Pearce [2]) that tests be simplified by estimating an analogue of (10) in two steps. It is proposed that both (5) and an analogous representation for $Y$ alone be estimated in a first step. Exogeneity can then be tested using statistics based on the cross-correlation functions of the residuals of the two estimated representations. The apparent simplifications available from this two-step procedure are, however, illusory.

In the initial step, we would estimate a system of the form

11) $$A(L)Y = C(L)U$$

$$H(L)X = J(L)W \quad ,$$

with $U$ and $W$ each a vector of white noises, $A$ and $H$ stable, $C$ and $J$ stable or borderline unstable, and identification achieved by some normalization rule. On the null hypothesis that $X$ is exogenous in a relation like (1) with $A$ in (1) invertible, lagged values of $U$ from (11) should be uncorrelated with current values of $W$, by the same argument which makes lagged values of $Y$ uncorrelated with current values of $W$ and justifies the test based on (10). Of course if the roles of $X$ and $Y$ are reversed, the possibility of $Y$'s being exogenous in a structural relation can be tested by examining whether lagged W's are correlated with current U's. This suggests a particularly simple and symmetrical procedure: find the cross-correlation of the U's with the W's at various relative lags, test first whether the cross-correlations are zero for all negative lags, then whether they are zero for all positive lags. This symmetrical procedure might look particularly attractive in bivariate systems where Y on X regressions and X on Y regressions are roughly equally plausible as structural relations.

The facts that U and W are each a vector of serially uncorrelated white noises, and that on the null hypothesis the particular set of cross-correlations being examined is zero, give the estimated cross-correlations of the true U and W a tractable asymptotic distribution. If $R_{UV}(t) = E(U(s)V(s-t)')$, if $R_{UV}(t) = 0$ for all $t$ greater than zero, and if $R^T_{UV}(t)$ is the estimate of $R_{UV}(t)$ formed from a sample on U and V of length T, then $T^{\frac{1}{2}}R^T_{UV}(t)$ converges in distribution to a matrix of independent normal variates. The joint distribution of a finite collection of such estimates at different positive values of t is also asymptotically normal, with the asymptotic covariances of estimates at different t zero.

However, in implementing this apparently natural procedure we must use estimates of U and W. Cross-covariances between U and W estimated from <u>estimated</u> U and W do not have the same asymptotic distribution as cross-covariances estimated from <u>actual</u> U and W, unless <u>all</u> cross-covariances between $U$ and $W$ are zero. In testing whether lagged values of $U$ are correlated with current $W$, we are in effect testing whether, in the equation

12) $$J^{-1}(L)H(L)X + K(L)U = W \quad ,$$

with $K(0)$ constrained to be zero, a non-zero $K$ contributes toward the explanation of current $X$. Replacing $U$ by $U^T$, the estimate of $U$ obtained from a sample of length $T$ on $Y$ and $X$, does not affect the legitimacy of a test based on joint estimates of $J^{-1}H$ and $K$. In fact, any weighted sum of current and past $Y$'s, including $Y$ itself as in (10), could replace $U$ in (12) without affecting the asymptotic validity of such a test. But replacing $W$ by $W^T$ and examining sample correlations as if they had been estimated from $W$ and $U$ instead of $W^T$ and $U^T$ amounts to estimating $J^{-1}H$ initially on the assumption that $K$ is zero, then examining the correlations of $U^T$'s with the $W^T$'s treating the estimated parameters of $J^{-1}H$ as fixed. When (12) is interpreted as a regression equation, it is easily understood that this leads to mistakes in inference. In a multivariate regression, correlations between estimated residuals and omitted variables whose true coefficients are zero have smaller asymptotic variance than correlations of the true residuals with those same omitted variables unless the omitted variables are uncorrelated with the included variables. This widely understood fact is a special case of a general result proved by Durbin [1]: If we have a vector of parameters $(a,b)$, if $a_0^T$ is the maximum likelihood estimate of $a$ conditional on $b=0$ from a sample of size $T$, and if $b_0^T$ is the maximum likelihood estimate of $b$ conditional on $a=a_0^T$ , then the asymptotic variance of $b_0^T$ on the null hypothesis that b=0 is smaller than the asymptotic variance which would be estimated on the assumption that $a_0^T$ is fixed, unless when the likelihood function is maximized jointly in a and b the resulting estimates are asymptotically uncorrelated. (For a more precise statement of this result, see the original Durbin article.) Thus this kind of procedure creates a bias, even in large samples, in favor of the null hypothesis $b=0$ . In our particular application, this is a bias in favor of $K=0$ , i.e. in favor of finding no feedback.

The bias vanishes only when U and X are uncorrelated at all lags.

To see how this bias works out in practice, consider this simple example:

13) $$Y(t) = X(t) + U(t)$$
$$X(t) = aX(t-1) + V(t) \quad .$$

Here X is in fact exogenous in the first equation of (13), under the usual assumptions that U and V are mutually and serially uncorrelated and $|a| < 1$. If we introduce $Y(t-1)$ on the right hand side of the second equation of (13), estimate a and the coefficient of $Y(t-1)$ jointly by OLS, and test the null hypothesis that the coefficient of lagged Y is zero by a standard t test, we have an asymptotically justified test of exogeneity. In fact, if we add any linear combination of past values of Y to the right-hand-side of (13) and test the coefficient of this additional variable by a t-test, we will have a legitimate test. But $Y(t-1)$ or the lagged value of any prewhitened version of

$Y$ will in general be correlated with $X(t-1)$. Estimating $a$ first in a regression of $X(t)$ on $X(t-1)$, then doing a separate standard test on the correlation of lagged $Y$ or lagged prewhitened $Y$ with the estimated residuals $\hat{V}$ from the first stage regression in effect ignores the effects of collinearity. In the limit as $a$ goes to zero and the variance of $U$ goes to zero, the bias in favor of the null hypothesis goes to infinity, in the sense that the true probability of type I error under (say) a nominal .05 level test goes to zero. This is large-sample bias--it does not diminish in large samples.

The bias of this two-step procedure is small if the relation between $Y$ and $X$ is weak. Thus if the null hypothesis to be tested is that $Y$ and $X$ are unrelated, tests based on the cross-correlogram of separately prewhitened series are legitimate. On the other hand, it is not safe to conclude that when the <u>statistical significance</u> of distributed lag relations between $Y$ and $X$ is low, tests based on the two-step procedure are not likely to be strongly biased.

For example, it would be natural to test for significance of the relation between $Y$ and $X$ by examing the sum of squares of the cross-correlations of pre-whitened versions of $Y$ and $X$. With, say, 100 quarters of data, we might examine lags in the range $\pm$ 12 quarters. To achieve significance at the ten per cent level under the usual asymptotic distribution, this sum of squares would have to exceed 34. But if the sum of squares of the true cross-correlations of whitened Y and whitened X were 34, the bias in the estimated variance of the correlation of a single lagged value of whitened Y with current whitened X could, on the null hypothesis that this correlation is zero, be as much as 34 per cent. This is sufficient bias to convert a nominal 5 per cent level test to 11 per cent, a nominal 10 per cent level test to 18 percent, a nominal 1 per cent level test to 4 per cent.

Of course, in addition there is the nearly obvious point that in many applications if we really believed that the relation between Y and X were very weak, there would be little point in checking carefully whether X is strictly exogenous in this very weak relation.

## Biblography


[1] Durbin, J., "Testing for Serial Correlation in Least Squares Regression when Some of the Regressors are Logged Dependent Variables", Econometrica, 38, p. 410-421, 1970.

[2] Feige, E., and K. D. Pearch, "The Causality Relationship Between Money and Income: A Time Series Approach", unpublished manuscript, 1974.

[3] Granger, C. W. J., "Investigating Causal Relations by Econometric Models and Cross-Sprectal Methods", Econometrica, 37, p. 424-238, 1969.

[4] Hannon, E. J., "Regression for Time Series" in Proceedings of a Symposium on Time Series Analysis, Mr. Rosenblatt (ed), (New York: John Wiley and Sons), 1963.

[5] Hannon, E. J., "The Identification problem for Multiple Equation Systems with Moving Average Errors", Econometrica, 39, p. 751-765, 1971.

[6] Phillips, A. W., "Some Notes on the Estimation of Time-Forms of Reactions in Interdependent Dynamic Systems", Econometrica, 23, p. 99-113, 1956.

[7] Pierce, D. A., and L. Haugh, "The Assessment and Detection of Causality in Temporal Systems", unpublished manuscript, Federal Reserve Board, May 1975.

[8] Rozanov, Yu. A., Stationary Random Processes (San Francisco: Holden-Day), 1967.

[9] Sims, C. A., "Money, Income and Causality", American Economic Review, 62, page 540-552, 1972.

[10] Zellner, A. and F. Palme, "Time Series Analysis and Simultaneous Equation Econometric Models", Journal of Econometrics, 2, p. 17-54, 1974.