Identifying Causal Relationships Between Nonstationary Stochastic Processes: An Examination of Alternative Approaches in Small Samples

Hector O. Zapata, Michael A. Hudson, and Philip Garcia

A Monte Carlo investigation is used to examine the performance of two commonly used tests for Granger causality for univariate and bivariate nonstationary *ARMA*(p,q) processes. Tests are applied to raw data, first differences of the raw data, and detrended versions of the series. The results indicate that for independent series the tests are robust regardless of sample size. With bivariate series and nonstationarity, the test results are sensitive to the *ARMA* specification, whether the data are filtered and the type of filter used, and the sample size.

Key words: causal relationships, Granger causality, lead/lag relationships, Monte Carlo, nonstationarity.

The concept of Granger testing has received considerable attention in recent years. Economists have found the approach particularly useful in analyzing temporal relationships between a variety of price series (e.g., Bessler and Brandt). Commonly, this approach has been used to identify lead/lag relationships between economic time series. Despite the widespread application of the technique, several questions continue regarding the method. Two primary issues have been raised with regard to the procedure: (*a*) what is the correct empirical approach, and (*b*) whether the tests are capable of correctly identifying causal relationships.

Various Monte Carlo studies have addressed these two issues. The general conclusions indicate the direct Granger method of testing, suggested by Sargent, outperforms the alternatives. This conclusion, however, is largely based on applications of the testing procedures to causally related stationary series which are primarily autoregressive in nature. It is important to examine the effectiveness of these tests with nonstationary data that follow more general autoregressive and moving-average *ARMA*(p,q), which are more common in economic time series. With nonstationary data, the application of the Granger test calls for tranformation of the data series to achieve stationarity. There is little empirical evidence regarding the performance of the Granger tests when nonstationary data series are used and stationarity inducing transformations employed prior to testing for causal relationships.

The purpose of this paper is to examine the performance of two tests of Granger causality for nonstationary *ARMA*(p,q) processes. Univariate and bivariate data series are constructed for alternative values of p and q. Different levels of covariance between the series are permitted, and various time trends which are typical of economic time series are added to induce nonstationarity. Two tests of Granger causality are then applied to the raw data, first differences of the raw data, and detrended versions of the series. Results suggest that for independent series the tests are robust. With bi-

The authors are, respectively, an assistant professor in the Department of Agricultural Economics and Agribusiness at Louisiana State University, and an assistant professor and an associate professor, both in the Department of Agricultural Economics at the University of Illinois.

Western Journal of Agricultural Economics, 13(2): 202–215

variate series and nonstationarity, the test results are influenced by the *ARMA* specification, the differences in nonstationarity, and the sample size.

The paper is organized as follows. Section two provides a brief summary of previous work. Experimental methods, including data generation procedures and application of the causality tests, are discussed in the third section. The fourth section of the paper presents the results of the inquiry. Last, the main implications and limitations of the study are summarized.

Background and Previous Work

The performance of various forms of the Granger test has been widely investigated. Monte Carlo studies by Nelson and Schwert; Geweke, Meese, and Dent; and Guilkey and Salemi are representative. In these studies, causally related stationary time series were constructed and a variety of tests applied. The general conclusions suggest the Granger tests are capable of identifying lead/lag relationships and that the direct Granger method yields the most consistent results. This method has also been widely used in applied studies between economic time series.

While these results verify the usefulness of these tests, the debate regarding the appropriate procedures for identifying causal relationships with nonstationary data continues. Lutkepohl concludes that "differencing nonstationary univariate component series of a multiple time series to induce stationarity prior to building an *AR* model for the multivariate generation process is in general inadequate" (p. 238). More recently, Bailey and Brorsen note, "there is no real agreement in the literature regarding the use or nonuse of a difference operator (prefilter) to obtain a stationary time series before causality tests are performed" (p. 128). Nerlove, Grether, and Carvalho, on the other hand, dismiss this issue, suggesting "letting the nonstationarity in one series explain the nonstationarity in the other" (p. 252).

Two recent empirical investigations are worthy of note with regard to the impacts of nonstationary series on the outcome of causality tests. Zeimer and Collins examined relationships between five agricultural price series and three theoretically unrelated series. The authors demonstrated that the Granger tests can identify relationships counter to theory when data possess nonstationary components. Bessler and Kling provide further evidence of the impact of nonstationary series in their investigation of sunspots (a stationary series) and gross national product (a nonstationary series). Using post-sample tests the authors demonstrate Granger tests give anomalous results when one series is stationary and the other nonstationary.

The results of these previous efforts suggest that the Granger approach is useful in identifying the lead/lag independence of stationary autoregessive time series but offer little insight into the effects of alternative generating processes on the performance of the tests. In particular, the effects of moving average components and nonstationary behavior which may be present in many economic series, deserve further attention and are considered below.

Experimental Methods

A Monte Carlo study was designed to examine the performance of two commonly used tests of Granger causality (the direct Granger and modified Sims) for *ARMA*(p,q) processes.[1] Table 1 summarizes the *ARMA*(p,q) data series used in the study, which were generated as stationary and nonstationary univariate and bivariate processes.[2] The univariate models were used as a benchmark for evaluation because they do not allow contemporaneous covariance between the error terms and have no constructed causal behavior ($C_{21} = 0$). The bi-

[1] In response to a reviewer's suggestion, nonstationary *ARIMA* $(p,1,q)$ processes were constructed for a supplemental investigation with 20 replications, using values of p and q between 0 and 2. These series provided an alternative form of nonstationarity for comparison with the time trends added to the *ARMA*(p,q) processes. The results of these models were not appreciably different from those presented below for the *ARMA*(p,q) models with nonstationarity in the form of time trends.

[2] Stationarity of process X_{t_0} implies that the characteristic polynomial (CP) $|I\epsilon^{p-1} - A_1\epsilon^{p-1} - \ldots - A_p\epsilon| \neq 0$ for $|\epsilon| \geq 1$. Alternatively, a p-dimensional difference equation with the CP equal to zero for which $|\epsilon\text{'s}| < 1$ provides stationary processes for all elements of A_k, $k = 1, 2, \ldots, p$. It can be shown that the coefficients in table 1 generate stationary processes by solving the CP for the difference equation of each model. For instance, for model $M2$ the solution to $|I\epsilon^2 - .5\epsilon - .25| = 0$ results in characteristic values ($\epsilon_1 = .81$, $\epsilon_2 = -.31$) whose absolute value is less than one. For the model $BM2$, the solution to $|I\epsilon^2 - A_1\epsilon - A_2| = 0$, where $A_1 = [(.5\ .0)'\ (.0\ .25)']$ and $A_2 = [(.6\ .0)'\ (.0\ .15)']$ has characteristic values ($\epsilon_1 = .81$, $\epsilon_2 = .79$, $\epsilon_3 = -.31$, and $\epsilon_4 = -.19$), which are also less than one in absolute value.

Table 1. *ARMA(p,q)* Processes Used to Generate Experimental Data

	AR Coefficient				MA Coefficient			
Model[a]	A_{11}	A_{12}	A_{21}	A_{22}	B_{11}	B_{12}	B_{21}	B_{22}
*M*1	.50	.00	.60	.00	.00	.00	.00	.00
*M*2	.50	.25	.60	.15	.00	.00	.00	.00
*M*3	.00	.00	.00	.00	.25	.00	.66	.00
*M*4	.00	.00	.00	.00	.25	.14	.66	.33
*M*5	.50	.00	.60	.00	.25	.00	.66	.00
*M*6	.50	.25	.60	.15	.25	.00	.66	.00
*M*7	.50	.25	.60	.15	.25	.14	.66	.33

Note: To induce nonstationary behavior, time-trend components were added to each of the above models. The following trend structures were imposed for the (*X*, *Y*) series: (0, 0), (.035, .035), and (.075, .035). Covariance between *X* and *Y* was analyzed at the .0, .1, .5, and .09, the causality parameter C_{21} equals .59. When referring to the corresponding bivariate model these acronyms are preceded by a *B*.
[a] The general form of the model is
$X(t) = A_{11}X(t-1) + A_{12}X(t-2) + B_{11}e(t-1) + B_{12}e(t-2) + C_{21}v(t-1) + e(t)$
$Y(t) = A_{21}Y(t-1) + A_{22}Y(t-2) + B_{21}v(t-1) + B_{22}v(t-2) + v(t)$.

variate processes were constructed under varying degrees (.1, .5, .9) of contemporaneous covariance between the terms $e(t)$ and $v(t)$ so as to analyze the impact on the causality tests. All the series (univariate and bivariate) were constructed with the following time trends for the (*X*, *Y*) pair: (.0, .0), (.035, .035), and (.075, .035).[3] Versions of each series were generated to examine the impacts of two sample sizes, 75 and 200 observations, on test performance.[4]

The stochastic components $e(t)$ and $v(t)$ were generated as independent normal (0, 1) variates and used directly to generate the univariate processes. The square root method (Rubinstein) was used to generate bivariate normal errors. For all series, the first fifty observations were deleted to minimize the effect of the starting values in the data generation process. Univariate and multivariate Lagrange multiplier tests (Jarque and Bera) were used to analyze the robustness of the methods in generating observations from the desired distribution. One hundred replications for the smallest sample size were run and the results indicated that at the .05 level of significance, we could be about 96% confident that $e(t)$ and $v(t)$ followed the specified distributions. The true model parameters could also be recovered with about the same level of significance.

The direct Granger and modified Sims tests as shown in table 2 were applied to each of the data series. Following applied procedures (e.g., Hsiao), univariate and multivariate versions of the Akaike's final prediction error (*FPE*) were used to determine the lag length for the series. A maximum lag length of six satisfied the criteria for most processes and was used throughout the evaluation. Both testing procedures were evaluated at the .05 level of significance.

Ashley; Ashley, Granger, and Schmalensee; Bessler and Kling; and others have recently demonstrated the benefits of using post-sample testing to verify causal relationships identified by Granger-type procedures. The primary motivation for post-sample testing is to verify that the causal relationships identified within the sample are not spurious. In cases where the true causal relationships are unknown, post-sample testing thus provides a means of verification of the identified causal relationships beyond the initial estimation period. Since the series being used in this investigation are generated by known processes, additional observations will follow a known process. Thus, post-sample tests would not be expected to alter the conclusions as the true

[3] As noted by an anonymous reviewer, these time trends reflect a fairly restrictive form of nonstationary behavior. However, many economic time series possess time trends. Although such behavior is clearly deterministic, conducting Granger causality tests on trended data can provide misleading results (Gamber and Hudson). The initial investigation reported in footnote 1 verifies tests performance for series constructed with other forms of nonstationary behavior.

[4] The entire simulation was performed on RATS. The simulation was very time consuming; it took approximately 18 hours to perform a complete analysis of one model, one sample size, 200 replications on an IBM-AT microcomputer with a math coprocessor. With 200 replications and a .05 level of significance, the results of the study will be accurate within a plus or minus 7.5% interval. In a power study, such as this, the error bands are derived from the variance of binominal distribution. Approximately 1,000 replications would be needed to obtain a plus or minus 4% interval. Computation time prohibited the use of a larger number of replications.

relationships are known a priori.[5] Nonetheless, in cases where the true model structures are unknown, such approaches are clearly useful and recommended to verify the results of Granger casusality tests.

Results

Causality test results for 75 and 200 observations are presented in tables 3 to 6. The three versions of each data series, raw, differenced, and detrended, are presented across the top of the table. Under each version of the data series, the direct Granger and modified Sims tests from table 2 are identified by $G2$, $G3$, $S1$, $S2$, and $S3$, respectively. Univariate models are represented by $M1$ to $M7$, and bivariate models by $BM1$ to $BM7$. The covariance level for the BM models is .5.[6] The three different time trends are represented by $T1$, $T2$, and $T3$ for (.0, .0), (.035, .035), and (.075, .035), respectively; a letter X or Y is added at the end of these time indices to define the dependent variable in the model of interest. Tables 3 and 4 contain the results for the univariate models. The values in the tables represent the proportion of rejections of the null hypothesis specified in table 2. If the tests are robust, the values should be close to zero with an error .05.

The results in table 3 indicate that both the direct Granger and modified Sims tests are robust in detecting lack of simultaneity and causal relationships when they did not exist. Except for a few cases, the percentage of rejections of the null hypothesis is very low. Increasing the model complexity from a purely autoregressive model of order one ($M1$) to an $ARMA(2,2)$ ($M7$) does not affect the test results at the .05 level of significance. In general, for univariate models, first differencing or detrending (using a linear time trend) does not have a major impact on the results.

Table 2. Test Specifications and Hypotheses

	Direct Granger Procedure
Model:	$x_t = \sum_{j=1}^{p} \theta_{1j} X_{t-j} + \sum_{i=0}^{q} \beta_{1i} Y_{t-i} + \epsilon_t$
$G2$:	H_0: $\beta_{10} = 0$, H_a: $\beta_{10} \neq 0$
$G3$:	H_0: $\beta_{11} = \ldots = \beta_{19} = 0$, H_a: Not all = 0
	Modified Sims Procedure
Model:	$X_t = \sum_{j=1}^{p} \theta_{2j} X_{t-j} + \sum_{k=1}^{r} \phi_{2k} Y_{t+k} + \sum_{i=0}^{q} \beta_{2i} Y_{t-i} + \nu_t$
$S1$:	H_0: $\phi_{21} = \ldots \phi_{2k} = 0$, H_a: Not all = 0
$S2$:	H_0: $\beta_{20} = 0$, H_a: $\beta_{20} \neq 0$
$S3$:	H_0: $\beta_{21} = \ldots = \beta_{2q} = 0$, H_a: Not all = 0

For the large sample size (200 observations, table 4) the introduction of time trend nonstationarity is of consequence, especially for the direct Granger ($G3$) test on raw data. However, the test performs well when first differencing or detrending are used to filter the data. This implies that either differencing or detrending are adequate filters when there is little or no contemporaneous covariance between the series.

Tables 5 and 6 provide results for bivariate models with causality constructed from Y and X and having .5 level of contemporaneous covariance. One would expect that if the test procedures are robust in detecting the true causal flows, the direct Granger ($G3$) and modified Sims ($S3$) tests should have values close to one when X is the dependent variable, and values close to zero when Y is the dependent variable. Similarly, test $S1$ should have values close to zero where X is the dependent variable and values close to zero when Y is the dependent variable. Because contemporaneous causation is tested by $G2$ and $S2$, their values should be close to one if the tests are robust.

For bivariate models, test results are influenced by nonstationarity, sample size and the $ARMA$ specification. The testing procedures for detecting simultaneous relationships produce somewhat mixed results. The direct Granger test ($G2$) provides very consistent results regardless of the model structure under consideration. Its accuracy increases with sample size and is invariant to the filtering pro-

[5] As noted by the editor and an anonymous reviewer, post-sample tests could be implemented in the context of the current study. For example, the forecasting performance of the estimated models could be compared with an alternative set of models generated using an alternative information criterion. Such an approach would provide verification of the ability of the tests to identify the true causal relationships between the data series. In light of the results presented below, which suggest the tests are indeed powerful in identifying the correct relationships, and the additional computational costs associated with these tests, no post-sample evaluation was performed.

[6] Preliminary results on 20 replications of the experiment indicated that for .1 level of covariance the results did not differ appreciably from those of univariate models, and that at high levels of covariance (.9) the results were not significantly different from those at .5 covariance.

Table 3. Monte Carlo Results for Univariate Models, 75 Observations, 200 Replications

Model		Raw Data					First Differences					Detrended				
		*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3
*M*1	*T*1*X*[a]	.040[b]	.025	.070	.055	.055	.070	.055	.055	.045	.040	.075	.040	.075	.055	.055
	*T*1*Y*	.040	.035	.060	.040	.040	.070	.025	.045	.050	.025	.075	.030	.050	050	.030
	*T*2*X*	.050	.155	.125	.035	.060	.050	.060	.070	.080	.055	.045	.050	.070	.045	.040
	*T*2*Y*	.050	.090	.095	.065	.020	.050	.030	.075	.050	.030	.045	.040	.070	.055	.015
	*T*3*X*	.080	.090	.070	.060	.065	.075	.070	.050	.095	.055	.065	.070	.055	.070	.065
	*T*3*Y*	.080	.150	.070	.075	.025	.075	.015	.045	.055	.020	.065	.020	.045	.070	.015
*M*2	*T*1*X*	.040	.045	.060	.060	.045	.065	.055	.055	.055	.040	.075	.070	.080	.060	.060
	*T*1*Y*	.040	.035	.045	.035	.035	.065	.020	.040	.055	.020	.075	.040	.045	.040	.020
	*T*2*X*	.050	.090	.095	.035	.065	.050	.045	.075	.090	.045	.045	.050	.075	.040	.045
	*T*2*Y*	.050	.045	.075	.050	.015	.050	.025	.065	.050	.025	.045	.040	.065	.055	.015
	*T*3*X*	.085	.070	.075	.050	.055	.060	.075	.060	.090	.055	.060	.080	.060	.060	.060
	*T*3*Y*	.085	.065	.070	.080	.025	.060	.020	.055	.075	.020	.060	.025	.050	.085	.015
*M*3	*T*1*X*	.050	.040	.065	.060	.055	.070	.065	.065	.035	.045	.075	.060	.065	.055	.050
	*T*1*Y*	.050	.025	.050	.035	.030	.070	.030	.040	.035	.030	.075	.025	.050	.045	.030
	*T*2*X*	.060	.175	.150	.055	.105	.050	.055	.075	.065	.060	.050	.045	.070	.055	.030
	*T*2*Y*	.060	.210	.090	.055	.035	.050	.025	.075	.040	.030	.050	.030	.060	.060	.015
	*T*3*X*	.070	.100	.090	.085	.090	.060	.070	.065	.100	.070	.060	.070	.040	.075	.060
	*T*3*Y*	.070	.305	.065	.085	.025	.060	.025	.045	.050	.015	.060	.020	.050	.075	.020
*M*4	*T*1*X*	.050	.030	.055	.065	.055	.065	.055	.060	.045	.055	.075	.060	.060	.070	.070
	*T*1*Y*	.050	.030	.050	.040	.025	.065	.025	.045	.030	.030	.075	.035	.055	.055	.025
	*T*2*X*	.060	.140	.135	.045	.085	.060	.065	.070	.070	.055	.045	.055	.070	.050	.040
	*T*2*Y*	.060	.160	.095	.070	.030	.060	.020	.075	.040	.030	.045	.035	.065	.065	.015
	*T*3*X*	.080	.090	.075	.070	.075	.070	.070	.060	.100	.060	.060	.070	.045	.075	.015
	*T*3*Y*	.080	.265	.065	.075	.025	.070	.015	.045	.050	.020	.060	.015	.050	.080	.015

Table 3. Continued

Model		Raw Data					First Differences					Detrended				
		G2	G3	S1	S2	S3	G2	G3	S1	S2	S3	G2	G3	S1	S2	S3
M5	*T1X*	.045	.040	.060	.035	.050	.065	.060	.060	.035	.045	.070	.075	.065	.045	.050
	T1Y	.045	.035	.050	.020	.030	.065	.030	.040	.055	.035	.070	.025	.040	.020	.030
	T2X	.050	.095	.115	.050	.045	.045	.055	.065	.045	.035	.045	.050	.080	.050	.040
	T2Y	.050	.055	.095	.055	.015	.045	.030	.065	.060	.030	.045	.040	.065	.065	.015
	T3X	.080	.085	.075	.075	.060	.065	.060	.050	.060	.060	.065	.065	.045	.065	.060
	T3Y	.080	.095	.060	.080	.020	.065	.020	.055	.090	.025	.065	.015	.050	.065	.010
M6	*T1X*	.045	.060	.070	.040	.050	.075	.055	.060	.045	.045	.065	.080	.090	.050	.060
	T1Y	.045	.035	.050	.025	.035	.075	.020	.040	.060	.025	.065	.035	.045	.020	.025
	T2X	.040	.050	.080	.045	.065	.045	.055	.070	.045	.025	.045	.055	.080	.040	.055
	T2Y	.040	.035	.065	.050	.020	.045	.035	.060	.045	.030	.045	.040	.060	.070	.010
	T3X	.070	.075	.070	.075	.050	.060	.060	.050	.045	.055	.060	.075	.050	.070	.065
	T3Y	.070	.060	.075	.075	.025	.060	.020	.060	.075	.020	.060	.025	.055	.070	.015
M7	*T1X*	.050	.055	.055	.055	.055	.060	.060	.060	.070	.055	.065	.070	.085	.055	.055
	T1Y	.050	.035	.055	.040	.035	.060	.025	.040	.055	.025	.065	.040	.050	.030	.025
	T2X	.035	.055	.085	.040	.055	.045	.060	.060	.045	.045	.045	.055	.080	.040	.055
	T2Y	.035	.040	.075	.060	.015	.045	.035	.060	.050	.015	.045	.040	.060	.060	.015
	T3X	.075	.055	.055	.060	.065	.060	.080	.035	.060	.065	.065	.070	.045	.065	.070
	T3Y	.075	.045	.070	.075	.020	.060	.020	.055	.075	.025	.065	.020	.065	.065	.015

[a] The models and tests are identified in tables 1 and 2, respectively. *T*1, *T*2, and *T*3 indicate the time trends used in constructing the data identified in table 1. The suffix immediately following the time trend indicator identifies the dependent variable.
[b] The numbers are the proportion of rejections of the null hypothesis for the tests identified in table 2 at the .05 level.

Table 4. Monte Carlo Results for Univariate Models, 200 Observations, 200 Replications

		Raw Data					First Differences					Detrended				
Model		*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3
*M*1	*T*1*X*	.040	.015	.040	.045	.020	.030	.020	.050	.040	.015	.040	.020	.040	.045	.025
	*T*1*Y*	.040	.000	.055	.060	.005	.030	.015	.070	.035	.005	.040	.000	.050	.055	.005
	*T*2*X*	.085	.360	.190	.045	.090	.070	.020	.045	.060	.040	.055	.015	.045	.060	.025
	*T*2*Y*	.085	.485	.125	.045	.065	.070	.005	.075	.055	.015	.055	.010	.070	.025	.010
	*T*3*X*	.090	.035	.085	.050	.020	.050	.005	.045	.040	.000	.065	.005	.030	.050	.005
	*T*3*Y*	.090	.890	.075	.055	.035	.050	.005	.070	.055	.005	.065	.000	.070	.055	.000
*M*2	*T*1*X*	.040	.020	.045	.040	.015	.040	.015	.040	.040	.020	.035	.025	.045	.040	.020
	*T*1*Y*	.040	.005	.060	.045	.000	.040	.005	.055	.040	.005	.035	.010	.055	.045	.000
	*T*2*X*	.075	.185	.105	.055	.055	.050	.025	.050	.055	.040	.050	.015	.050	.060	.015
	*T*2*Y*	.075	.230	.085	.040	.030	.050	.010	.080	.005	.010	.050	.010	.065	.030	.015
	*T*3*X*	.065	.025	.065	.050	.000	.050	.000	.040	.040	.000	.050	.005	.030	.050	.010
	*T*3*Y*	.065	.545	.060	.040	.015	.050	.000	.075	.045	.000	.050	.000	.070	.050	.005
*M*3	*T*1*X*	.040	.015	.040	.045	.015	.035	.010	.060	.035	.010	.040	.020	.045	.055	.015
	*T*1*Y*	.040	.005	.055	.050	.000	.035	.015	.070	.030	.005	.040	.005	.060	.050	.005
	*T*2*X*	.145	.400	.235	.065	.180	.075	.010	.045	.060	.030	.060	.020	.045	.075	.020
	*T*2*Y*	.145	.790	.165	.065	.200	.075	.015	.070	.065	.025	.060	.010	.065	.045	.010
	*T*3*X*	.090	.040	.100	.045	.025	.055	.010	.045	.045	.010	.060	.000	.040	.050	.005
	*T*3*Y*	.090	.990	.100	.065	.075	.055	.005	.065	.070	.000	.060	.000	.060	.065	.000
*M*4	*T*1*X*	.040	.015	.045	.045	.020	.030	.010	.055	.040	.010	.040	.020	.035	.055	.020
	*T*1*Y*	.040	.000	.055	.045	.000	.030	.010	.070	.040	.005	.040	.000	.055	.040	.000
	*T*2*X*	.125	.320	.220	.060	.145	.075	.025	.045	.065	.030	.060	.020	.040	.065	.015
	*T*2*Y*	.125	.755	.145	.070	.160	.075	.005	.075	.070	.025	.060	.010	.070	.045	.010
	*T*3*X*	.085	.040	.090	.040	.020	.050	.005	.050	.050	.000	.065	.000	.030	.045	.005
	*T*3*Y*	.085	.980	.080	.050	.065	.050	.005	.070	.075	.000	.065	.000	.060	.070	.000

Table 4. Continued

Model		Raw Data					First Differences					Detrended				
		*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3
*M*5	*T*1*X*	.035	.015	.045	.055	.020	.030	.015	.055	.040	.005	.025	.020	.045	.055	.020
	*T*1*Y*	.035	.000	.055	.060	.005	.030	.005	.065	.020	.000	.025	.000	.065	.060	.005
	*T*2*X*	.080	.130	.165	.055	.095	.075	.020	.050	.080	.035	.065	.020	.060	.045	.025
	*T*2*Y*	.080	.590	.100	.040	.045	.075	.005	.075	.060	.045	.065	.010	.070	.035	.010
	*T*3*X*	.050	.015	.065	.030	.010	.045	.000	.045	.045	.000	.055	.000	.030	.035	.005
	*T*3*Y*	.050	.840	.065	.070	.010	.045	.005	.075	.055	.000	.055	.000	.065	.060	.000
*M*6	*T*1*X*	.040	.015	.040	.055	.015	.030	.015	.055	.060	.015	.035	.020	.045	.055	.015
	*T*1*Y*	.040	.005	.055	.055	.005	.030	.005	.055	.030	.000	.035	.000	.060	.055	.005
	*T*2*X*	.065	.090	.100	.050	.055	.055	.020	.045	.070	.030	.055	.020	.055	.050	.020
	*T*2*Y*	.065	.275	.080	.035	.030	.055	.010	.075	.050	.010	.055	.010	.060	.035	.015
	*T*3*X*	.050	.010	.050	.025	.000	.035	.000	.035	.040	.000	.055	.005	.030	.035	.010
	*T*3*Y*	.050	.490	.050	.065	.010	.035	.000	.060	.045	.000	.055	.005	.060	.065	.005
*M*7	*T*1*X*	.045	.015	.045	.060	.015	.035	.015	.045	.050	.020	.040	.020	.045	.055	.015
	*T*1*Y*	.045	.005	.060	.035	.005	.035	.005	.045	.035	.005	.040	.010	.060	.035	.005
	*T*2*X*	.060	.055	.080	.055	.055	.055	.020	.045	.055	.015	.045	.020	.055	.050	.020
	*T*2*Y*	.060	.200	.075	.025	.030	.055	.005	.085	.060	.005	.045	.010	.065	.035	.010
	*T*3*X*	.045	.010	.050	.040	.000	.040	.005	.030	.050	.005	.055	.005	.030	.040	.010
	*T*3*Y*	.045	.405	.055	.060	.015	.040	.000	.075	.040	.000	.055	.000	.060	.045	.000

Note: See table 3 for a description of the table.

Table 5. Monte Carlo Results for Bivariate Models, 75 Observations, 200 Replications

Model		Raw Data					First Differences					Detrended				
		G2	*G3*	*S1*	*S2*	*S3*	*G2*	*G3*	*S1*	*S2*	*S3*	*G2*	*G3*	*S1*	*S2*	*S3*
BM1	*T1X*	.980	.690	.050	.945	.315	.935	.695	.065	.955	.905	.975	.675	.075	.935	.320
	T1Y	.980	.025	.950	.335	.040	.975	.020	.940	1.000	.150	.975	.030	.945	.040	.035
	T2X	.995	.845	.080	.960	.250	.950	.685	.080	.945	.905	.980	.695	.090	.960	.235
	T2Y	.995	.010	.945	.395	.025	.990	.020	.910	1.000	.140	.980	.015	.920	.400	.025
	T3X	1.000	.745	.090	.930	.385	.945	.740	.045	.945	.920	.980	.750	.060	.930	.345
	T3Y	1.000	.100	.940	.320	.100	.985	.010	.935	.985	.150	.980	.010	.940	.370	.070
BM2	*T1X*	.980	.700	.050	.935	.480	.945	.675	.060	.965	.920	.975	.680	.070	.935	.410
	T1Y	.980	.040	.945	.405	.340	.970	.025	.940	1.000	.085	.975	.025	.940	.495	.170
	T2X	.985	.800	.070	.945	.395	.965	.675	.065	.950	.935	.985	.695	.075	.955	.330
	T2Y	.985	.010	.930	.430	.330	.990	.020	.910	1.000	.060	.985	.015	.910	.495	.180
	T3X	.985	.765	.065	.930	.520	.955	.740	.040	.955	.935	.980	.780	.045	.945	.435
	T3Y	.985	.060	.940	.385	.370	.980	.015	.935	.995	.080	.980	.010	.945	.418	.215
BM3	*T1X*	.980	.755	.045	.900	.195	.935	.760	.075	.890	.785	.980	.745	.065	.915	.220
	T1Y	.980	.025	.980	.640	.020	.975	.020	.960	.990	.290	.980	.025	.980	.695	.005
	T2X	1.000	.885	.085	.960	.140	.950	.800	.065	.920	.785	.980	.810	.080	.925	.165
	T2Y	1.000	.025	.965	.750	.000	.990	.015	.955	1.000	.255	.980	.015	.955	.700	.000
	T3X	.995	.823	.105	.895	.205	.945	.820	.040	.935	.800	.980	.825	.060	.915	.185
	T3Y	.995	.200	.960	.575	.030	.985	.005	.955	.990	.230	.980	.005	.955	.660	.020
BM4	*T1X*	.980	.750	.060	.915	.360	.940	.750	.070	.955	.880	.975	.725	.080	.915	.340
	T1Y	.980	.020	.975	.725	.025	.965	.025	.955	.990	.260	.975	.030	.975	.765	.030
	T2X	1.000	.855	.085	.970	.365	.965	.790	.070	.960	.900	.980	.785	.080	.945	.350
	T2Y	1.000	.040	.970	.790	.005	.990	.015	.955	1.000	.215	.980	.010	.955	.750	.005
	T3X	1.000	.830	.090	.890	.485	.950	.815	.045	.945	.875	.980	.820	.055	.905	.425
	T3Y	1.000	.185	.960	.600	.085	.975	.010	.960	.995	.190	.980	.010	.955	.710	.035

Table 5. Continued

Model		Raw Data					First Differences					Detrended				
		G2	*G3*	*S1*	*S2*	*S3*	*G2*	*G3*	*S1*	*S2*	*S3*	*G2*	*G3*	*S1*	*S2*	*S3*
BM5	*T1X*	.980	.770	.040	.645	.400	.955	.745	.065	.910	.385	.975	.760	.060	.655	.315
	T1Y	.980	.040	.980	.070	.050	.980	.010	.955	.950	.100	.975	.050	.980	.095	.045
	T2X	.995	.845	.070	.675	.345	.965	.805	.080	.935	.360	.980	.830	.075	.695	.300
	T2Y	.995	.040	.960	.055	.025	.990	.020	.950	.935	.085	.980	.015	.955	.075	.025
	T3X	.995	.810	.055	.685	.645	.955	.815	.050	.925	.345	.980	.820	.055	.695	.375
	T3Y	.995	.065	.955	.100	.120	.980	.005	.955	.915	.080	.980	.010	.965	.145	.060
BM6	*T1X*	.980	.076	.045	.660	.600	.960	.740	.070	.945	.390	.975	.765	.055	.670	.475
	T1Y	.980	.040	.980	.130	.300	.980	.020	.955	.975	.040	.975	.030	.980	.150	.175
	T2X	.995	.850	.060	.700	.630	.970	.810	.070	.950	.390	.985	.825	.090	.720	.495
	T2Y	.995	.015	.965	.105	.175	.990	.015	.950	.975	.015	.985	.010	.090	.155	.145
	T3X	.990	.805	.055	.715	.785	.955	.805	.060	.935	.415	.980	.830	.050	.735	.530
	T3Y	.990	.030	.960	.155	.345	.975	.005	.955	.960	.040	.980	.010	.965	.215	.225
BM7	*T1X*	.985	.760	.055	.665	.760	.930	.730	.075	.950	.525	.975	.745	.070	.700	.695
	T1Y	.985	.035	.980	.150	.435	.980	.020	.960	.975	.030	.975	.040	.985	.170	.320
	T2X	.995	.810	.065	.745	.790	.935	.780	.085	.970	.620	.980	.795	.075	.770	.705
	T2Y	.995	.015	.970	.140	.305	.985	.010	.950	.980	.005	.980	.010	.965	.190	.270
	T3X	.985	.805	.055	.735	.875	.925	.810	.055	.955	.580	.980	.835	.055	.765	.765
	T3Y	.985	.030	.955	.190	.470	.975	.015	.960	.970	.035	.980	.020	.960	.230	.305

Note: See table 3 for a description of the table.

Table 6. Monte Carlo Results for Bivariate Models, 200 Observations, 200 Replications

Model		Raw Data					First Differences					Detrended				
		*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3
*BM*1	*T*1*X*	1.000	1.000	.045	1.000	.900	1.000	1.000	.040	1.000	1.000	1.000	1.000	.050	1.000	.885
	*T*1*Y*	1.000	.010	1.000	.815	.280	1.000	.025	1.000	1.000	.670	1.000	.015	1.000	.825	.270
	*T*2*X*	1.000	1.000	.075	1.000	.855	1.000	1.000	.055	1.000	1.000	1.000	1.000	.055	1.000	.865
	*T*2*Y*	1.000	.015	1.000	.840	.245	1.000	.020	1.000	1.000	.665	1.000	.005	1.000	.830	.215
	*T*3*X*	1.000	.995	.120	1.000	.905	1.000	1.000	.045	1.000	1.000	1.000	.995	.020	1.000	.800
	*T*3*Y*	1.000	.785	1.000	.810	.535	1.000	.000	1.000	1.000	.690	1.000	.015	1.000	.880	.220
*BM*2	*T*1*X*	1.000	1.000	.035	1.000	.940	1.000	1.000	.050	1.000	1.000	1.000	1.000	.055	1.000	.785
	*T*1*Y*	1.000	.015	1.000	.875	.985	1.000	.020	1.000	1.000	.300	1.000	.020	1.000	.885	.965
	*T*2*X*	1.000	1.000	.060	1.000	.985	1.000	1.000	.055	1.000	1.000	1.000	1.000	.045	1.000	.980
	*T*2*Y*	1.000	.020	1.000	.860	.955	1.000	.025	1.000	1.000	.370	1.000	.010	1.000	.895	.915
	*T*3*X*	1.000	1.000	.080	1.000	.985	1.000	1.000	.040	1.000	1.000	1.000	.995	.020	1.000	.940
	*T*3*Y*	1.000	.335	1.000	.910	.985	1.000	.005	1.000	1.000	.370	1.000	.010	1.000	.915	.905
*BM*3	*T*1*X*	1.000	1.000	.040	1.000	.510	1.000	1.000	.065	1.000	1.000	1.000	1.000	.045	1.000	.495
	*T*1*Y*	1.000	.010	1.000	.990	.015	1.000	.015	1.000	1.000	.950	1.000	.015	1.000	.995	.025
	*T*2*X*	1.000	1.000	.125	1.000	.575	1.000	1.000	.060	1.000	1.000	1.000	1.000	.060	1.000	.590
	*T*2*Y*	1.000	1.000	1.000	.990	.020	1.000	.025	1.000	1.000	.905	1.000	.015	1.000	.990	.020
	*T*3*X*	1.000	1.000	.140	1.000	.590	1.000	1.000	.055	1.000	1.000	1.000	1.000	.030	1.000	.495
	*T*3*Y*	1.000	.960	1.000	.990	.090	1.000	.000	1.000	1.000	.915	1.000	.015	1.000	.995	.015
*BM*4	*T*1*X*	1.000	1.000	.035	1.000	.950	1.000	1.000	.045	1.000	1.000	1.000	1.000	.040	1.000	.950
	*T*1*Y*	1.000	.010	1.000	1.000	.100	1.000	.010	1.000	1.000	.885	1.000	.010	1.000	1.000	.085
	*T*2*X*	1.000	1.000	.160	1.000	.945	1.000	1.000	.055	1.000	1.000	1.000	1.000	.050	1.000	.945
	*T*2*Y*	1.000	.175	1.000	1.000	.065	1.000	.020	1.000	1.000	.825	1.000	.005	1.000	1.000	.075
	*T*3*X*	1.000	1.000	.115	1.000	.965	1.000	1.000	.045	1.000	1.000	1.000	1.000	.020	1.000	.730
	*T*3*Y*	1.000	.950	1.000	.990	.370	1.000	.005	1.000	1.000	.840	1.000	.015	1.000	.995	.065

Table 6. Continued

Model		Raw Data					First Differences					Detrended				
		*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3	*G*2	*G*3	*S*1	*S*2	*S*3
*BM*5	*T*1*X*	1.000	1.000	.050	.995	.980	1.000	1.000	.065	1.000	.910	1.000	1.000	.050	.990	.980
	*T*1*Y*	1.000	.010	1.000	.175	.255	1.000	.015	1.000	1.000	.390	1.000	.015	1.000	.185	.240
	*T*2*X*	1.000	1.000	.130	.985	.975	1.000	1.000	.055	1.000	.925	1.000	1.000	.065	.980	.940
	*T*2*Y*	1.000	.170	1.000	.175	.220	1.000	.010	1.000	1.000	.430	1.000	.015	1.000	.175	.225
	*T*3*X*	1.000	1.000	.090	.985	1.000	1.000	1.000	.055	1.000	.915	1.000	1.000	.025	.990	.955
	*T*3*Y*	1.000	.835	1.000	.130	.590	1.000	.010	1.000	1.000	.405	1.000	.010	1.000	.225	.210
*BM*6	*T*1*X*	1.000	1.000	.055	.995	1.000	1.000	1.000	.060	1.000	.920	1.000	1.000	.055	.995	1.000
	*T*1*Y*	1.000	.010	1.000	.320	.955	1.000	.015	1.000	1.000	.040	1.000	.010	1.000	.335	.925
	*T*2*X*	1.000	1.000	.070	.985	.995	1.000	1.000	.055	1.000	.965	1.000	1.000	.055	.985	.995
	*T*2*Y*	1.000	.040	1.000	.310	.875	1.000	.010	1.000	1.000	.060	1.000	.015	1.000	.355	.870
	*T*3*X*	1.000	1.000	.070	.990	1.000	1.000	1.000	.035	1.000	.940	1.000	1.000	.030	.995	.985
	*T*3*Y*	1.000	.485	1.000	.285	.975	1.000	.015	1.000	1.000	.060	1.000	.020	1.000	.345	.860
*BM*7	*T*1*X*	1.000	1.000	.050	.995	1.000	1.000	1.000	.050	1.000	.985	1.000	1.000	.055	.995	1.000
	*T*1*Y*	1.000	.015	1.000	.385	.990	1.000	.015	1.000	1.000	.035	1.000	.020	1.000	.400	.985
	*T*2*X*	1.000	1.000	.075	.995	1.000	1.000	1.000	.050	1.000	.990	1.000	1.000	.040	.995	1.000
	*T*2*Y*	1.000	.035	1.000	.380	.960	1.000	.010	1.000	1.000	.035	1.000	.010	1.000	.415	.955
	*T*3*X*	1.000	1.000	.065	.990	1.000	1.000	1.000	.040	1.000	.990	1.000	1.000	.030	.995	1.000
	*T*3*Y*	1.000	.395	1.000	.390	1.000	1.000	.015	1.000	1.000	.045	1.000	.010	1.000	.435	.970

Note: See table 3 for a description of the table.

cedure used. In contrast, the modified Sims approach ($S2$) for detecting simultaneity does not perform well. While its accuracy improves with sample size, its performance is influenced by model specification, nonstationarity, and the method of filtering. Generally, its accuracy improves when the nonstationarity is removed with the first difference filter.

For the lead/lag relationships, again the results are mixed. The modified Sims test for examining lags ($S3$) is heavily influenced by model specification, type of nonstationarity, and filtering procedure. Its accuracy does not improve with sample size nor with any particular filtering procedure. In fact, filtering often influences this test in an unpredictable manner, calling into question its use in applied work. The results are somewhat different for the direct Granger test for lags ($G3$) and the modified Sims test for leads ($S1$). In the raw data form, the results of the $G3$ test are influenced rather dramatically by the type of nonstationarity—its accuracy declining as the nonstationarity takes different forms (i.e., different time trends). Filtering the data by either first differences or detrending improves the accuracy of the test. In large, filtered samples, the $G3$ test provides the most accurate identification of lead/lag relationships. The $S1$ test is very accurate in small and large samples. It is least affected by nonstationarity, method of filtering and sample size. For smaller samples, it provides the most accurate procedure for identifying lead/lag relationships. Its accuracy declines marginally with large samples in the raw data form, but filtering improves its performance to an acceptable level.

Summary, Implications, and Limitations

The results of the Monte Carlo analysis into the performance of the Granger and modified Sims tests on the presence of nonstationarity have important implications for agricultural economists interested in identifying lead/lag relationships between economic time series. Economic data possess nonstationary components and model identification is not always clearly specified by theory. The findings of the current study provide insight into the usefulness of procedures often applied.

For univariate *ARMA* processes, the direct Granger and modified Sims tests are robust in detecting lack of simultaneity and unidirectional causal relationships. Differencing or detrending improves the test performance for nonstationary series. Hence, when the degree of contemporaneous covariance between two stochastic processes is very low, the tests provide reliable results.

For bivariate processes, on balance, the results suggest that simultaneous relationships can be best identified by the Granger test ($G2$). Lead/lag relationships in the raw data form are most accurately identified using the modified Sims ($S1$) procedure. When nonstationarity is removed by filtering the data, $G3$ performs best in large samples while $S1$ is the most accurate with smaller samples. Hence, for smaller series generated by bivariate processes, perhaps the situation which most closely approximates many analyses of economic time series, the $S1$ test should be used to complement the more widely used $G3$ test to establish lead/lag relationships.

In general, the results of the simulation indicate that the effects of nonstationarity are most pronounced in larger samples which possess dissimilar trends, and for tests examining lagged variables. On the whole, except for the case of the modified Sims test ($S3$), first differencing of the bivariate processes does not significantly disturb the bivariate nature of the processes and permits identification of the correct lead/lag relationships.

There are several limitations which should be noted. While a rather wide range of models and trends was selected for analysis for the study, the conclusions of the research are strictly applicable only to the models and forms of nonstationarity tested. The results, therefore, cannot be generalized to series which possess other forms of deterministic behavior, such as seasonal or cyclic components. Also, the study considered only first difference transformations and detrending procedures to induce stationarity and therefore says nothing about the performance when other transformations (e.g., *ARIMA* filters) are performed. In addition, the sensitivity of the results to alternative selection criteria was not examined. Nevertheless, the *FPE* did seem to provide model structures which permitted consistent identification of the underlying causal structures.

Despite these limitations, the results should prove useful for researchers interested in examining lead/lag relationships between economic time series. They corroborate previous findings that Granger tests and selected modified Sims tests can be used with confidence in applied work in identifying various bivariate relationships. While nonstationarity complicates their application, judicious filtering or selection of appropriate tests can remedy some of these concerns. Further work needs to be done to identify how tests perform when other deterministic behavior exists in the data and the series generated by more complicated model structures.

[Received March 1986; final revision received June 1988.]

References

Ashley, R. "Inflation and the Distribution of Price Changes Across Markets: A Causal Analysis." *Econ. Inquiry* 19(1981):650–60.

Ashley, R., C. W. J. Granger, and R. Schmalensee. "Advertising and Consumption: An Analysis of Causality." *Econometrica* 48(1980):1149–68.

Bailey, D., and B. W. Brorsen. "Dynamics of Regional Fed Cattle Prices." *West. J. Agr. Econ.* 10(1984): 126–33.

Bessler, D. A., and J. A. Brandt. "Causality Tests in Livestock Markets." *Amer. J. Agr. Econ.* 64(1982):140–44.

Bessler, D. A., and J. L. Kling. "A Note on Tests of Granger Causality." *Appl. Econ.* 16(1984):335–42.

Gamber, E. N., and M. A. Hudson. "Causality Testing with Messy Data: Some Preliminary Experimental Evidence." Selected paper, presented at AAEA meeting, Ithaca, NY, Aug. 1984. (Abstract in *Amer. J. Agr. Econ.* 66(1984):900.)

Geweke, J., R. Meese, and W. T. Dent. "Comparing Alternative Tests of Causality in Temporal Systems: Analytic Results and Experimental Evidence." *J. Econometrics* 21(1983):161–94.

Granger, C. W. J. "Investigating Causal Relations by Econometric Models and Cross Spectral Methods." *Econometrica* 37(1969):424–38.

Guilkey, D. K., and M. K. Salemi. "Small Sample Properties of Three Tests for Granger Causal Ordering in a Bivariate Stochastic System." *Rev. Econ. and Statist.* 64(1982):668–80.

Hsiao, C. "Autoregressive Modeling of Canadian Money and Income Data." *J. Amer. Statist. Assoc.* 74(1979): 554–60.

Jarque, C. M., and A. K. Bera. "Efficient Tests for Normality, Homoskedasticity, and Serial Independence of Regression Residuals." *Econ. Letters* 6(1980):255–59.

Lutkepohl, H. "Differencing Multiple Time Series: Another Look at Canadian Money and Income Data." *J. Time Ser. Anal.* 3(1982):235–43.

Nelson, C. R., and G. W. Schwert. "Tests for Predictive Relationships between Time-Series Variables: A Monte Carlo Investigation." *J. Amer. Statist. Assoc.* 37(1982):11–18.

Nerlove, M., D. M. Grether, and J. L. Carvalho. *Analysis of Economic Time Series.* New York: Academic Press, 1979.

Rubinstein, R. Y. *Simulation and the Monte Carlo Method.* New York: John Wiley & Sons, 1981.

Sargent, T. J. "A Classical Macroeconomic Model for the United States." *J. Polit. Econ.* 84(1976):207–38.

Zeimer, R. F., and G. S. Collins. "Granger Causality and U.S. Crop and Livestock Prices." *S. J. Agr. Econ.* 16(1984):115–19.