Mixed Unit Roots and Deterministic Trends in Noncausality Tests

Tao Ran

Hector Zapata*

Louisiana State University

Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meeting, Dallas, TX, February 2-6, 2008

*Tao Ran is a Ph.d candidate, Hector Zapata is a professor at the department of Agricultural Economics, LSU. The corresponding author is Tao Ran, with e-mail tran1@lsu.edu, and telephone (225)578-2595

Mixed Unit Roots and Deterministic Trends in Noncausality Tests

Tao Ran

Hector Zapata

Abstract Using Japanese economic data and a Monte Carlo simulation, this study analyzes the consequences of ignoring deterministic trends in mixed unit-root data for Granger noncausality tests. Results from an augmented VAR suggest over-rejection in certain empirically relevant cases at various sample sizes.

1. Introduction

In time series studies, time can be incorporated as an explanatory variable to model a deterministic trend. In spite of the role played by the deterministic trend term, some empirical studies in economics often proceed by ignoring the trend term, that is, not incorporating time into the model as an explanatory variable, even when the presence of a deterministic trend is suspected. This might not be a problem in some cases, but in others it can cause a distortion in the statistical hypothesis test which might lead to a wrong conclusion. This study uses a Monte Carlo simulation experiment to analyze the consequence of ignoring the trend term in a lag augmented vector autoregressive (VAR) model using a modified Wald test to test for Granger noncausality. The results show that ignoring a deterministic trend in a modified Wald test rejects a true null hypothesis more often than it should be.

2. Literature Review

2.1. Trend

In time series analysis, "the key feature of a trend is that it has a permanent effect on a series" (Enders 2004). If a model consists of a trend term and a stationary[1] component, the sequence will exhibit only temporary departures from the trend. This type of model is defined as trend stationary. A linear trend term in a model as an explanatory variable is called deterministic trend. Now, if the error term, or shock as it is called in time series, has a permanent effect on changing the conditional mean of the series, the

[1] Or, covariance stationary. A time series is covariance stationary if its mean and all autocovariances are unaffected by a change of time origin, or simply independent of time. For more details, see Enders 2004, page 52.

sequence is said to have a stochastic trend. A simple example of a stochastic trend model is the random walk model which has the simple expression of $y_t = y_{t-1} + \varepsilon_t$, where ε_t is white noise[2]. When a process has a deterministic trend, we need to take this into account. However, in some time series analyses, people do not make a distinction between the model having a trend and the one that does not. Oftentimes, a universal test was used whether a deterministic trend existed or not. This study evaluates the impact of ignoring a deterministic trend when testing Granger noncausality in time series analysis.

2.2 Causality

The majority of time series causality analyses are based on the idea of Granger causality. Basically, it is not possible for a cause to come after the effect. If, for example, a variable z is affected by a variable x, then the prediction of z should be improved by including x in the information set. According to Lütkepohl (2005), suppose Ω_t is the information set containing all of the relevant information in the universe available up to and including period t. Let $z_t(h|\Omega_t)$[3] be the optimal (minimum mean square error or MSE) h-step predictor of the process z_t at origin t based on the information in Ω_t. The corresponding forecast MSE will be denoted by $\Sigma_z(h|\Omega_t)$. The process x_t is said to cause z_t in Granger's sense if

$$\Sigma_z(h|\Omega_t) < \Sigma_z(h|\Omega_t \backslash \{x_s \mid s \leq t\}) \quad \text{for some } h = 1, 2, \ldots , \tag{1}$$

where $\Omega_t \backslash \{x_s \mid s \leq t\}$ is the set containing all of the relevant information in the universe except for the information in the past and present of the x_t process. In other words, the left hand side of the inequality is the MSE based on the whole information set, especially information about the past and present values of x_t, while the right hand side does not include the whole information set, where \ means exclusion.

If the above holds, x_t is said to Granger-cause z_t or x_t is Granger-causal for z_t. Alternatively, if z_t can be predicted more efficiently when the information in the x_t

---

[2] A sequence $\{\varepsilon_t\}$ is called a white-noise process if each value in the sequence has a mean of zero, a constant variance, and is uncorrelated with all other realizations (Enders 2004).

[3] For example, when h = 1, the expression $z_t(1|\Omega_t)$ means the 1 step predictor of the process z_t based on the information up to and including period t, that is, predicting z_{t+1} using all of the information available at period t.

process is taken into account in addition to all other information in the universe, then x_t is Granger-causal for z_t. Note that this definition can be extended to the case where z and x are more than a one dimensional process.

The Granger noncausality test is often used in the economics literature to detect the relationship between certain economic variables. For example, there has been a debate on the relationship between exports and a country's economic growth during past decades. Some argue that export is the 'engine' of economic growth (export-led growth hypothesis, ELG hypothesis hereafter) (Keesing 1967, Krueger 1985). Others say that the former was just a 'handmaiden' (growth-led exports hypothesis, GLE hypothesis hereafter) (Bhagwati 1988, Lancaster 1980, Krugman 1984). But the empirical evidence continues to be mixed. One driver of these mixed results could be related to the misspecification of deterministic time trends.

2.3 Granger noncausality Test Approaches

In applied work, three approaches are frequently used to investigate Granger noncausality: formal tests of restrictions, impulse response functions (IRF), and forecast error variance decompositions (FEVD). This study is focused on the formal restriction test based on the theoretical framework discussed in the above section. In the context of the ELG, Granger non-causality is tested via a likelihood ratio, Wald, or F test to check the validity of the exclusion restrictions. Often, however, in analyzing economic variables, a VAR process may have non-stationary elements such as unit roots and/or cointegration[4]. This will "alter the asymptotic distributional results of the least square (LS) estimators of the coefficients which results in Granger non-causality test statistics that may not have standard asymptotic null distributions" (Giles and Williams 2000).

A vector error correction model (VECM) is usually used to test for causality when cointegration is suspected. However, when the process has nonstationary structure or a mixed-unit-root structure, the testing of these restrictions is not as straightforward as that for a stationary process. This may lead to over-rejection of the non-causality null and/or wrong conclusions of causality (Giles and Williams 2000b). Further, it requires a pretest of unit roots and cointegration which might be sensitive to lag orders. Another issue in

[4] In economic analysis, series with unit roots is a special case of the non-stationary process. Cointegration is a phenomenon whereby two or more series with unit roots may be related. Detailed discussion is in Appendix B.6.

using VECM is the specification of the model structure. A common practice in many empirical analyses is not to distinguish different model structures via deterministic trends.

An alternative approach by Toda and Yamamoto (1995) and Dolado and Lütkepohl (1996) (TYDL hereafter), which is the focus of this study, does not require a pretest of unit root and cointegration, and thus can lead to asymptotically equivalent results. Usually a Wald test is chosen to test for non-causality, which has an asymptotic χ^2 distribution if the process is stationary. However, if the system is non-stationary, the asymptotic covariance matrix of the coefficient estimator is singular. TYDL proposed a lag augmented VAR to solve the problem. The basic idea, as discussed in LÜtkepohl (2005), is that whenever the elements in at least one of the complete coefficient matrices A_i are not restricted under H_0, the Wald statistic has its usual asymptotic χ^2 distribution. In other words, if restrictions are placed on all A_i, i = 1, … , p, as in the non-causality hypothesis, we can get a χ^2 Wald test by adding an extra lag in estimating the parameters of the process. The good news is that we do not have to know the cointegration properties of the process to use this testing method. The bad news, however, is that there will be a loss of power in this regard. If the dimension K is much smaller than the number of lags p, the loss will be small though. Further, if the system is bivariate and cointegrated of rank one with both variables being I(1), then no extra lag is needed.

The test is called a modified Wald test, and it is easy to apply in empirical research. The merit of the modified Wald test is that it does not require a pretest of unit roots and/or cointegration. Also, it accommodates cases where some series in the system are I(1) and others are I(0), that is, when a mixed order of integration is found for the individual series. The test is not affected by the order of integration, but it is affected by the existence of a deterministic trend term. Unfortunately, in practice, people often ignore the trend term and use with the model structure without a trend. An illustration of the difference the trend term makes is presented in the following section using Japan's economy data as an example, followed by a Monte Carlo experiment concerning this matter.

3. Empirical Analysis

3.1 Data

We use quarterly data on Japan during the period of 1960-2006 to test the ELG hypothesis which is, in essence, a Wald test of Granger noncausality. Five variables were chosen to be included in the analysis: GDP, export, terms of trade (export divided by import), gross fixed capital formation as a proxy for capital investment, and an industrial production index for all the industrialized countries as a proxy for foreign output shock[5].

Although the lag augmented VAR approach does not require a test of unit roots, for the sake of the specific question of interest, we would like to know whether the individual series are I(1) or I(0) with or without a deterministic trend. In the literature, two tests are common to test the existence of unit roots for each individual variable, namely, the Dickey-Fuller test and the Philips-Perron test. The former is sensitive to the lag order of the variables, while the latter provides an alternative for the heterogeneous or weakly dependent unit root process. Plus, the Philips-Perron test does not have reliable results for small sample sizes. The results of the tests are in the tables below.

Table1. Unit root test results for exports.

Statistic	Dickey-Fuller	p-value (D-F)	Phillips-Perron	p-value (P-P)	Conclusion
τ	1.71	.98	3.51	.99	I(1) with
τ_μ	-3.37	.01	-3.37	.01	trend
Φ_1	8.12	.001			
τ_τ	-2.35	.41	-1.64	.77	
Φ_3	5.91	.07			

Table2. Unit root test results for terms of trade

Statistic	Dickey-Fuller	p-value (D-F)	Phillips-Perron	p-value (P-P)	Conclusion
τ	-1.51	.12	-2.27	.02	I(0)
τ_μ	-2.59	.097	-3.08	.03	
Φ_1	3.68	.13			
τ_τ	-2.32	.42	-3.95	.01	
Φ_3	3.72	.43			

Table3. Unit root test results for GDP

Statistic	Dickey-Fuller	p-value (D-F)	Phillips-Perron	p-value (P-P)	Conclusion
τ	-.21	.61	3.97	.99	I(0)
τ_μ	-2.78	.06	-7.63	.001	
Φ_1	3.88	.096			
τ_τ	-1.42	.85	-.50	.98	
Φ_3	.28	.28			

[5] The choice of these variables is based on economic theory and previous literature.

Table4. Unit root test results for industrial production index

Statistic	Dickey-Fuller	p-value (D-F)	Phillips-Perron	p-value (P-P)	Conclusion
τ	2.3	.99	4.13	.99	I(1) with
τ_μ	-2.62	.09	-2.77	.07	trend
Φ_1	6.78	.001			
τ_τ	-3.03	.13	-3.08	.12	
Φ_3	6.37	.049			

Table5. Unit root test results for capital investment

Statistic	Dickey-Fuller	p-value (D-F)	Phillips-Perron	p-value (P-P)	Conclusion
τ	1.77	.98	4.32	.99	I(1) with
τ_μ	-3.23	.02	-6.56	.001	trend
Φ_1	7.37	.001			
τ_τ	- .80	.96	-1.28	.89	
Φ_3	6.17	.06			

From the results, we conclude that export is I(1) with a deterministic trend, terms of trade is I(0) without trend, GDP is I(0) without trend, industrial production is I(1) with deterministic trend, and capital investment is I(1) with deterministic trend. Because we are interested in whether Japan's export is leading the country's economic growth or the other way around, GDP and exports are the two main variables of concern. Note that the lag order[6] for the VAR model is chosen by the corrected Akaike information criterion (AIC) when there is no trend in it and the lag order of six has the minimum value. For the augmented model, a lag of seven is chosen to get the residual error to calculate the modified Wald test. When there is a trend, three criteria were used for the lag order selection: AIC, HQ and SC[7]. The lag order of five has the minimum value for two of the criteria.

3.2 Basic model

The basic model structure is[8]

---

[6] The selection of the lag order is to specify the value of p in VAR(p) which is similar to model selection in regression.

[7] AIC stands for Akaike's Information Criterion, HQ stands for the Hannan-Quinn Criterion, and SC is the Schwarz Criterion using Bayesian arguments. Details are in Lütkepohl (2005). The order selection is done using PROC IML in SAS, with codes attached in Appendix C.4.

[8] For the model without trend terms, another lag, that is, lag seven is going to be added for the augmented model. For the model with trend, a trend term will be added.

$$\begin{bmatrix} ext_t \\ gdp_t \\ tot_t \\ inv_t \\ prd_t \end{bmatrix} = \mathrm{V} + \mathrm{A}_1 \times \begin{bmatrix} ext_{t-1} \\ gdp_{t-1} \\ tot_{t-1} \\ inv_{t-1} \\ prd_{t-1} \end{bmatrix} \ldots \ldots + \mathrm{A}_6 \times \begin{bmatrix} ext_{t-6} \\ gdp_{t-6} \\ tot_{t-6} \\ inv_{t-6} \\ prd_{-6t} \end{bmatrix} + \varepsilon_t ,$$

where V is the vector of intercepts and

$$\mathrm{A}_i = \begin{bmatrix} a_{11,i} & \cdots & a_{15,i} \\ a_{21,i} & \cdots & a_{25,i} \\ & \cdots & \\ a_{51,i} & \cdots & a_{55,i} \end{bmatrix}$$

for all i = 1, 2, …, 6. Here ext stands for export, gdp stands for gross domestic product, tot stands for terms of trade, inv stands for capital investment, and prd stands for industrial production. To test for the GLE hypothesis, we would like to test whether $a_{12,i}$ is zero for all i. If we can not reject that they are all zero, growth is not Granger causing export. Similarly, to see whether the ELG hypothesis is true, we would check whether $a_{21,i}$ is zero for all i. If we reject the null that they are all zero, export is Granger causing growth in Japan during that period of time. We use the seemingly unrelated regression (SUR) model[9] to estimate this VAR(6) of dimension five.

Table 6 gives the results of the modified Wald test in two cases: the one with deterministic trend terms included in the export, investment, and industrial production equations, and the one ignoring the deterministic terms. From the results, we can see that ignoring the deterministic trend term affects the modified Wald test result, but just to a small degree. Both models conclude that there is no causal effect between export and growth in Japan, that is, export is not led by economic growth, and economic growth is not led by export.

The slight difference between the two might be due to the data used and/or the true model structure, which is never known by the researcher. To see generally how ignoring the deterministic trend would affect the rejection accuracy of the modified Wald

[9] It is a set of simultaneous equation model with each equation having different coefficients for the same explanatory variables. The model also allows for different explanatory variables. The equations may appear uncorrelated, but the correlation across the errors in different equations can provide links that can be exploited in estimation. Details about using SUR to calculate a Wald test are in Rambaldi (1996).

test, a Monte Carlo experiment was conducted, the merit of which is that the data generating process (DGP) is known and can be controlled.

Table 6.Wald test results for models with and without a trend term

Model with deterministic trend terms			Model ignoring deterministic trend		
Hypothesis	Wald test	p-value	Hypothesis	Wald test	p-value
ELG	5.729	.333	ELG	1.934	.858
GLE	8.809	.117	GLE	8.46	.132

3.3 Monte Carlo Experiment

For our purpose, we generated 1000 samples of a system of time series of either I(1) or I(0) with deterministic trend in some of the series. Also, one series is simulated to Granger cause another. Hence, we know the true process has a trend term in it and that the two series have a causal relationship between them. Then, we set up a SUR model discussed in the previous section, but using the generated data instead. The series in the SUR model ignores the deterministic trend term on purpose. A modified Wald test is calculated based on the SUR model estimation. Because we know the true model has trend but the estimation model ignores the trend, and we know exactly which series is Granger causing another, we can tell how much distortion ignoring the trend term causes in detecting the causal relationship.

The criteria in designing the GDP are model dimension, direction of causality, stability, and error structure. Bivariate processes are included first in the experiment due to its simplicity and their frequent appearance in economic empirical work. Higher dimensional models are considered also. However, in a Monte Carlo experiment, higher dimensional models are hard to manage and to interpret, so we include models with at most three variables. In all models there is a causality from series a to series b in the true process. There are mainly three scenarios to be considered. For example, in the bivariate case, we have both series are I(1) and one has a deterministic trend term; both are I(1) and both have trend terms; and one is I(1) without trend and the other is I(0) with trend. The trivariate case is similar but adds one more series c, which is I(1) with a trend term.

For each scenario, we have one lag, two lags, and four lags specified in the true model as three different cases to accommodate the real life data that are either yearly or quarterly. Also, three error structure forms are considered: identity covariance matrix, which is the simplest case; heterogeneous but contemporaneously uncorrelated variance-covariance matrix; and the most general one, heterogeneous and contemporaneously correlated variance-covariance matrix. Further, because in real practice the true lag order is never known and has to be selected using certain criteria, we might have the selected lag order less than the true order (short lag) or more than the true order (long lag). Usually the short lag is one lag less than the true lag and the long lag is one lag more than the true. This is also taken into consideration in the experiment. In all cases 1000 samples of size T + 50 are generated with the first 50 observations discarded[10]. For each DGP, five sample sizes were included: T = 25, 50, 100, 200, and 300.

The tabulated results of the experiment are listed in Tables 7 and 8. Table 7 contains the outcome for the bivariate models, while Table 8 presents the results for the trivariate models. Recall that the true causality is a is causing b, so we wanted to reject the null of a not causing b often (close to 100%), and reject the null of b not causing rarely (close to 5%). In all cases the numbers in the body of the tables are the percentage of rejections at the 5 percent significance level.

From Table 7 we can see that ignoring the deterministic trend term is not problematic in detecting the true direction of causality. As the right block of Table 7 indicates, when the true relationship is "a causes b," we do not have much trouble rejecting the null that a is not causing b, especially when the sample size is large. This is true for all the scenarios. However, ignoring the deterministic trend will cause over-rejection of the hypothesis that is true, as shown in the left block of the table. We know b is not causing a in the true DGP, and expect the null of b is not causing a to be "accepted" as often as possible. Unfortunately, if we do not specify the trend term in the estimating model, we will have lower power for the modified Wald test. The problem worsens when both "a and b" have a trend term which are ignored by the estimated model. As shown in scenario 2 of the table, the null of b not causing a is often rejected even though it is true.

[10] This is to eliminate the influence of the different initial values of the sequence.

Moreover, increasing the sample size will not correct the problem but make it worse. The problem is less obvious when only one series has trend, whether it is I(0) or I(1).

If the true error structure is heterogeneous with a correlated variance-covariance, more over-rejection will occur. This makes sense because the Wald test based on a lag augmented VAR assumes a white noise structure for the error term. If the lag order in the true model is four, the problem is a little bit worse overall. In general but not always, selecting a short lag will have less over-rejection than finding the true lag or selecting a long lag. The trivariate case has similar results except that when one more series having a trend is included in the model that are composed of two series with trend also, the over-rejection problem is less serious than the bivariate one. This is from the comparison of scenario 2 in the bivariate and trivariate cases. However, for scenario 1 and 3, adding one more series with trend will cause more over-rejection

Table 7. Percent Rejection of Non-causality Using Modified Wald test, 5 percent level, 1000 samples (two series)

Scenario 1	Lag Order	Error Structure	Lag Selection	H_0: b is NOT causing a Sample size					H_0: a is NOT causing b Sample size				
				25	50	100	200	300	25	50	100	200	300
Both series are I(1) and one of the series has trend	1 LAG	IDENTITY	TRUE	0.092	0.088	0.079	0.1	0.308	0.675	0.923	1	1	1
			LONG	0.12	0.101	0.105	0.094	0.396	0.98	0.999	1	1	1
		HETEROGENEOUS	TRUE	0.068	0.05	0.052	0.057	0.061	0.89	0.923	0.973	1	1
			LONG	0.085	0.06	0.057	0.057	0.06	0.972	0.996	1	1	1
		MOST GENERAL	TRUE	0.203	0.148	0.414	0.954	0.998	0.884	0.902	0.952	1	1
			LONG	0.207	0.174	0.573	0.995	1	0.982	0.996	1	1	1
	2 LAGS	IDENTITY	SHORT	0.061	0.067	0.084	0.109	0.134	0.81	0.99	1	1	1
			TRUE	0.074	0.071	0.07	0.065	0.089	0.991	1	1	1	1
			LONG	0.095	0.06	0.069	0.059	0.072	0.985	1	1	1	1
		HETEROGENEOUS	SHORT	0.066	0.059	0.077	0.065	0.065	0.829	0.919	0.981	1	1
			TRUE	0.081	0.055	0.065	0.041	0.057	0.965	0.998	1	1	1
			LONG	0.101	0.063	0.064	0.052	0.056	0.966	0.996	1	1	1
		MOST GENERAL	SHORT	0.159	0.216	0.38	0.66	0.814	0.853	0.936	0.981	0.999	1
			TRUE	0.18	0.252	0.419	0.744	0.903	0.966	0.995	1	1	1
			LONG	0.161	0.194	0.329	0.654	0.857	0.971	0.995	1	1	1
	4 LAGS	IDENTITY	SHORT	0.102	0.05	0.054	0.044	0.035	0.975	0.999	1	1	1
			TRUE	0.125	0.057	0.052	0.041	0.041	0.965	0.999	1	1	1
			LONG	0.153	0.068	0.054	0.039	0.041	0.942	0.999	1	1	1
		HETEROGENEOUS	SHORT	0.101	0.053	0.068	0.069	0.076	0.996	1	1	1	1
			TRUE	0.115	0.061	0.073	0.051	0.063	0.997	1	1	1	1
			LONG	0.145	0.056	0.058	0.05	0.05	0.992	1	1	1	1
		MOST GENERAL	SHORT	0.134	0.094	0.109	0.201	0.28	0.943	0.988	1	1	1
			TRUE	0.16	0.098	0.141	0.3	0.455	0.941	0.989	0.999	1	1
			LONG	0.195	0.114	0.154	0.289	0.462	0.939	0.992	1	1	1

Table 7---continued

Scenario 2	Lag Order	Error Structure	Lag Selection	H_0: b is NOT causing a — Sample size 25	50	100	200	300	H_0: a is NOT causing b — Sample size 25	50	100	200	300
Both series are I(1) and both have trend	1 LAG	IDENTITY	TRUE	0.105	0.075	0.205	0.995	1	0.665	0.898	0.996	1	1
			LONG	0.129	0.103	0.3	0.99	1	0.985	0.997	1	1	1
		HETEROGENEOUS	TRUE	0.071	0.049	0.046	0.116	0.416	0.87	0.907	0.965	0.999	1
			LONG	0.084	0.061	0.053	0.159	0.45	0.975	0.995	1	1	1
		MOST GENERAL	TRUE	0.15	0.147	0.401	0.951	0.995	0.869	0.899	0.939	0.991	1
			LONG	0.204	0.175	0.568	0.998	1	0.982	0.995	1	1	1
	2 LAGS	IDENTITY	SHORT	0.117	0.179	0.332	0.637	0.786	0.608	0.915	0.999	1	1
			TRUE	0.126	0.177	0.298	0.588	0.789	0.97	1	1	1	1
			LONG	0.116	0.139	0.224	0.495	0.721	0.978	1	1	1	1
		HETEROGENEOUS	SHORT	0.082	0.085	0.137	0.236	0.325	0.759	0.855	0.925	0.981	0.997
			TRUE	0.072	0.078	0.124	0.192	0.314	0.803	0.956	0.998	1	1
			LONG	0.088	0.075	0.084	0.127	0.178	0.944	0.99	1	1	1
		MOST GENERAL	SHORT	0.245	0.42	0.706	0.953	0.996	0.729	0.807	0.901	0.973	0.995
			TRUE	0.261	0.463	0.755	0.987	0.998	0.94	0.987	1	1	1
			LONG	0.219	0.356	0.668	0.961	0.999	0.953	0.992	1	1	1
	4 LAGS	IDENTITY	SHORT	0.384	0.785	0.998	1	1	0.834	0.986	1	1	1
			TRUE	0.4	0.709	0.989	1	1	0.724	0.929	1	1	1
			LONG	0.311	0.408	0.823	0.998	1	0.736	0.928	1	1	1
		HETEROGENEOUS	SHORT	0.239	0.611	0.978	1	1	0.966	1	1	1	1
			TRUE	0.25	0.571	0.983	1	1	0.924	0.999	1	1	1
			LONG	0.218	0.298	0.785	0.999	1	0.938	1	1	1	1
		MOST GENERAL	SHORT	0.33	0.756	0.996	1	1	0.82	0.848	0.905	1	0.996
			TRUE	0.325	0.711	0.988	1	1	0.833	0.836	0.906	1	0.999
			LONG	0.354	0.588	0.929	1	1	0.838	0.805	0.873	1	0.97

Table 7---continued

Scenario 3	Lag Order	Error Structure	Lag Selection	H_0: b is NOT causing a — Sample size					H_0: a is NOT causing b — Sample size				
				25	50	100	200	300	25	50	100	200	300
One series is I(0) with trend and the other is I(1) without trend	1 LAG	IDENTITY	TRUE	0.11	0.089	0.087	0.093	0.091	0.329	0.618	0.923	0.998	1
			LONG	0.13	0.118	0.105	0.095	0.092	0.363	0.692	0.973	1	1
		HETEROGENEOUS	TRUE	0.078	0.058	0.053	0.048	0.048	0.833	0.83	0.888	0.982	0.986
			LONG	0.096	0.069	0.053	0.059	0.063	0.81	0.818	0.911	0.992	1
		MOST GENERAL	TRUE	0.161	0.175	0.397	0.695	0.841	0.848	0.871	0.921	0.975	0.99
			LONG	0.212	0.215	0.428	0.711	0.849	0.847	0.871	0.953	0.994	0.999
	2 LAGS	IDENTITY	SHORT	0.046	0.043	0.084	0.092	0.14	0.765	0.98	1	1	1
			TRUE	0.073	0.054	0.062	0.076	0.094	0.967	1	1	1	1
			LONG	0.094	0.057	0.063	0.056	0.081	0.948	1	1	1	1
		HETEROGENEOUS	SHORT	0.059	0.087	0.13	0.154	0.215	0.838	0.926	0.987	0.999	1
			TRUE	0.071	0.062	0.073	0.076	0.097	0.935	0.993	1	1	1
			LONG	0.085	0.068	0.067	0.058	0.069	0.927	0.985	1	1	1
		MOST GENERAL	SHORT	0.101	0.135	0.271	0.514	0.699	0.856	0.953	0.997	1	1
			TRUE	0.191	0.309	0.55	0.908	0.983	0.965	0.998	1	1	1
			LONG	0.186	0.244	0.458	0.815	0.964	0.934	0.996	1	1	1
	4 LAGS	IDENTITY	SHORT	0.108	0.059	0.055	0.064	0.06	0.891	0.999	1	1	1
			TRUE	0.12	0.068	0.05	0.052	0.055	0.851	0.995	1	1	1
			LONG	0.137	0.062	0.056	0.053	0.052	0.818	0.989	1	1	1
		HETEROGENEOUS	SHORT	0.073	0.056	0.057	0.057	0.06	0.996	1	1	1	1
			TRUE	0.107	0.055	0.053	0.045	0.067	0.995	1	1	1	1
			LONG	0.126	0.057	0.05	0.055	0.049	0.993	1	1	1	1
		MOST GENERAL	SHORT	0.141	0.102	0.175	0.421	0.637	0.852	0.969	0.999	1	1
			TRUE	0.14	0.122	0.198	0.519	0.704	0.879	0.968	0.997	1	1
			LONG	0.181	0.135	0.206	0.459	0.66	0.871	0.966	0.996	1	1

Table 8. Percent Rejection of Non-causality Using Modified Wald test, 5 percent level, 1000 samples (three series)

Scenario 1	Lag Order	Error Structure	Lag Selection	H_0: b is NOT causing a Sample size					H_0: a is NOT causing b Sample size				
				25	50	100	200	300	25	50	100	200	300
Both series are I(1) and one of the series has trend	1 LAG	IDENTITY	TRUE	0.067	0.06	0.051	0.041	0.054	0.589	0.906	0.996	1	1
			LONG	0.077	0.068	0.056	0.054	0.055	0.527	0.88	0.999	1	1
		HETEROGENEOUS	TRUE	0.085	0.063	0.056	0.04	0.322	0	0	0	0.07	0.898
			LONG	0.115	0.072	0.06	0.05	0.304	0.435	0.386	0.419	0.344	0.048
		MOST GENERAL	TRUE	0.399	0.385	0.34	0.274	0.32	0.923	0.925	0.968	0.985	0.99
			LONG	0.48	0.454	0.409	0.486	0.378	0.921	0.933	0.977	0.988	0.994
	2 LAGS	IDENTITY	SHORT	0.076	0.074	0.059	0.053	0.062	0.703	0.972	1	1	1
			TRUE	0.102	0.066	0.044	0.046	0.06	0.868	0.999	1	1	1
			LONG	0.126	0.087	0.048	0.054	0.058	0.863	0.999	1	1	1
		HETEROGENEOUS	SHORT	0.077	0.055	0.062	0.072	0.076	0.419	0.999	1	1	1
			TRUE	0.089	0.068	0.048	0.053	0.067	0.995	1	1	1	1
			LONG	0.121	0.082	0.055	0.053	0.059	0.992	1	1	1	1
		MOST GENERAL	SHORT	0.268	0.243	0.253	0.249	0.263	0.871	0.894	0.962	0.986	0.996
			TRUE	0.386	0.348	0.323	0.352	0.345	0.896	0.933	0.983	0.998	1
			LONG	0.434	0.335	0.314	0.344	0.322	0.905	0.937	0.984	0.999	1
	4 LAGS	IDENTITY	SHORT	0.135	0.084	0.068	0.066	0.061	0.857	0.999	1	1	1
			TRUE	0.183	0.087	0.057	0.049	0.058	0.811	0.995	1	1	1
			LONG	0.288	0.081	0.058	0.035	0.061	0.782	0.977	1	1	1
		HETEROGENEOUS	SHORT	0.142	0.072	0.063	0.065	0.069	0.982	1	1	1	1
			TRUE	0.169	0.091	0.07	0.054	0.063	0.977	0.999	1	1	1
			LONG	0.277	0.076	0.073	0.05	0.056	0.962	0.999	1	1	1
		MOST GENERAL	SHORT	0.405	0.348	0.339	0.394	0.408	0.883	0.964	0.999	1	1
			TRUE	0.486	0.3888	0.324	0.369	0.372	0.914	0.953	0.998	1	1
			LONG	0.582	0.351	0.337	0.348	0.324	0.938	0.94	0.99	1	1

Table 8---continued

Scenario 2	Lag Order	Error Structure	Lag Selection	H_0: b is NOT causing a — Sample size 25	50	100	200	300	H_0: a is NOT causing b — Sample size 25	50	100	200	300
Both series are I(1) and both have trend	1 LAG	IDENTITY	TRUE	0.065	0.065	0.053	0.052	0.085	0.549	0.883	0.995	1	1
			LONG	0.079	0.063	0.064	0.062	0.07	0.487	0.835	0.998	1	1
		HETEROGENEOUS	TRUE	0.109	0.078	0.067	0.048	0.341	0	0	0	0.21	0.986
			LONG	0.125	0.078	0.066	0.057	0.345	0.389	0.327	0.377	0.426	0.144
		MOST GENERAL	TRUE	0.394	0.356	0.319	0.273	0.307	0.917	0.924	0.971	0.994	0.995
			LONG	0.482	0.436	0.401	0.379	0.369	0.919	0.936	0.98	0.992	0.999
	2 LAGS	IDENTITY	SHORT	0.085	0.073	0.066	0.067	0.086	0.687	0.965	1	1	1
			TRUE	0.104	0.064	0.055	0.062	0.094	0.839	0.999	1	1	1
			LONG	0.133	0.076	0.048	0.067	0.083	0.825	0.998	1	1	1
		HETEROGENEOUS	SHORT	0.107	0.097	0.179	0.323	0.473	0.579	0.994	1	1	1
			TRUE	0.177	0.187	0.314	0.062	0.783	0.554	0.741	1	1	1
			LONG	0.203	0.153	0.249	0.463	0.658	0.796	0.989	1	1	1
		MOST GENERAL	SHORT	0.259	0.238	0.252	0.244	0.255	0.876	0.91	0.97	0.996	0.998
			TRUE	0.382	0.344	0.319	0.344	0.336	0.904	0.94	0.988	1	1
			LONG	0.43	0.33	0.309	0.337	0.317	0.895	0.945	0.993	1	1
	4 LAGS	IDENTITY	SHORT	0.157	0.095	0.118	0.169	0.22	0.773	0.996	1	1	1
			TRUE	0.21	0.097	0.097	0.109	0.136	0.712	0.975	1	1	1
			LONG	0.299	0.089	0.087	0.083	0.11	0.722	0.939	1	1	1
		HETEROGENEOUS	SHORT	0.124	0.063	0.078	0.127	0.17	0.895	0.992	1	1	1
			TRUE	0.173	0.086	0.081	0.084	0.12	0.875	0.987	1	1	1
			LONG	0.283	0.076	0.07	0.073	0.095	0.9	0.987	1	1	1
		MOST GENERAL	SHORT	0.357	0.305	0.29	0.329	0.358	0.922	0.986	1	1	1
			TRUE	0.457	0.358	0.3	0.352	0.331	0.943	0.977	0.999	1	1
			LONG	0.572	0.343	0.311	0.341	0.318	0.954	0.974	0.998	1	1

Table 8---continued

Scenario 3	Lag Order	Error Structure	Lag Selection	H_0: b is NOT causing a Sample size 25	50	100	200	300	H_0: a is NOT causing b Sample size 25	50	100	200	300
One series is I(0) with trend and the other is I(1) without trend	1 LAG	IDENTITY	TRUE	0.068	0.068	0.047	0.042	0.051	0.597	0.909	0.996	1	1
			LONG	0.086	0.077	0.056	0.055	0.052	0.489	0.824	0.993	1	1
		HETEROGENEOUS	TRUE	0.068	0.068	0.047	0.042	0.05	0.597	0.909	0.996	1	1
			LONG	0.105	0.06	0.061	0.05	0.051	0.41	0.343	0.348	0.474	0.633
		MOST GENERAL	TRUE	0.308	0.246	0.283	0.257	0.271	0.862	0.902	0.953	0.98	0.986
			LONG	0.422	0.345	0.346	0.345	0.35	0.873	0.898	0.947	0.974	0.987
	2 LAGS	IDENTITY	SHORT	0.079	0.078	0.049	0.038	0.061	0.667	0.97	1	1	1
			TRUE	0.111	0.067	0.051	0.052	0.054	0.776	0.996	1	1	1
			LONG	0.114	0.075	0.062	0.05	0.057	0.703	0.984	1	1	1
		HETEROGENEOUS	SHORT	0.077	0.058	0.059	0.048	0.047	0.165	0.936	1	1	1
			TRUE	0.093	0.067	0.043	0.045	0.052	0.996	1	1	1	1
			LONG	0.127	0.078	0.051	0.056	0.07	0.985	1	1	1	1
		MOST GENERAL	SHORT	0.259	0.239	0.232	0.237	0.243	0.855	0.892	0.961	0.988	0.997
			TRUE	0.392	0.341	0.338	0.336	0.34	0.867	0.907	0.964	0.995	1
			LONG	0.42	0.319	0.278	0.322	0.317	0.89	0.908	0.952	0.983	0.996
	4 LAGS	IDENTITY	SHORT	0.142	0.075	0.074	0.06	0.062	0.697	0.986	1	1	1
			TRUE	0.204	0.087	0.058	0.049	0.059	0.642	0.95	1	1	1
			LONG	0.265	0.09	0.056	0.049	0.066	0.67	0.911	1	1	1
		HETEROGENEOUS	SHORT	0.135	0.049	0.05	0.066	0.063	0.983	1	1	1	1
			TRUE	0.176	0.085	0.59	0.046	0.058	0.973	1	1	1	1
			LONG	0.265	0.083	0.05	0.049	0.049	0.95	1	1	1	1
		MOST GENERAL	SHORT	0.342	0.316	0.299	0.353	0.394	0.853	0.904	0.978	0.998	1
			TRUE	0.444	0.374	0.325	0.354	0.349	0.883	0.9	0.955	0.988	0.999
			LONG	0.552	0.352	0.319	0.329	0.329	0.907	0.901	0.948	0.988	0.996

4. Summary and Conclusion

This study focuses on the difference that a deterministic trend term makes in detecting a causal relationship between two series. A Granger noncausality test is commonly used to identify the relation between series. As an alternative to the VECM structure, modified Wald test is more general and flexible. Further, it allows for mixed unit roots in the VAR model. Often in empirical work, however, models are specified using lagged explanatory variables and a Wald test applied. This simple approach may not always work well. Using a specific real life data set and a Monte Carlo experiment, we find that ignoring the trend term causes over-rejection of a true hypothesis using a modified Wald test. In other words, the test can mislead people in concluding that causality exists between certain series when there is none in reality, due to not incorporating the deterministic trend term. The problem is more serious in the bivariate case when the errors of the two series have heterogeneous variances and are contemporaneously correlated. Moreover, the over-rejection problem cannot be corrected when the sample size gets larger, say over 200. As indicated in scenario 2 of Table 7 and 8, when the sample size is big and the error structure is most general, over 95 percent of the time we are going to reject the true null. In summary, care should be taken in constructing time series models that have deterministic trend, especially in a VAR mixed unit roots context. A parsimonious lagged variable specification might draw the modeled structure too far away from the true data generating process. Phillips and Ploberger (1996) and Phillips (1996) discuss in detail the criteria for choosing the representations for trending time series.

Recent work by Phillips (2002) showed that a stochastic trend can be explained by deterministic functions and/or lagged variables. There is indeed a competition between the two. If the modeling of a stochastic trend by deterministic function is designed seriously and carefully, it will be very close to the true model. On the other hand, from the parsimony and forecasting perspectives, lagged variable models are simpler. The best way is to carefully inspect the data and design the trend model such that it is parsimonious yet still capture the essential features of the trending behavior. If we simply use the lagged terms, it might be parsimonious, but some characteristics of the trending behavior is neglected and one of the consequences is the inaccuracy of the relevant tests.

Robinson and Iacone (2005) develop the properties of estimates of the cointegrating coefficient in a bivariate model that either ignore or incorporate additive deterministic trends.

There are at least two aspects that need to be further analyzed in the future work of this paper. First, a comparison of the cointegrating test performance with VECM specification and TYDL Wald test should be done using Monte Carlo experiment in a mixed unit root context. It would be of theoretical and practical interest to assess Phillips (2002) findings in the context of mixed unit roots and expand this work by comparing VECMs with augmented VARs as done in the Monte Carlo simulation in this study. Further work is also needed in post-sample evaluations of alternative specifications of mixed unit roots and deterministic trends as suggested by Phillips (2002). The findings in this study provide initial evidence that much caution is needed in the specification of deterministic trends for samples of size commonly found in empirical work when some series are I(1) and others are I(0). One approach that holds significant merits in this context, yet for the most part has been ignored in agricultural economic applications, is the adoption of fully modified multivariate regressions. Work in progress addresses such issues.

References

Awokuse T. O. (2007), "Export-led Growth and the Japanese Economy: Evidence from VAR and Directed Acyclic Graphs," *Applied Economics*, 38:5, 593-602

Bhagwati, J.N. (1988), *Protectionism,* Cambridge, MA: MIT Press

Boltho, A. (1996), "Was Japanese growth export-led?," *Oxford Economic Papers*, 48, 415–32

Dolado J.J. and H. Lutkepohl (1996), "Making Wald Tests Work for Cointegrated VAR Systems" *Econometric Reviews*, 15(4), 369-386 (1996)

Davidson, R and J. G. MacKinnon (2004), *Economic Theory and Methods*, New York, Oxford University Press, 2004

Engle, R. and Granger, C.(1987). *Co-Integration and Error Correction: Representation, Estimation, and Testing* Econometrica, Vol. 55, No. 2. (Mar., 1987), pp. 251-276.

Enders, W. (1995), *Applied Econometric Time Series*, John and Wiley Sons. 1995.

Giles, J A. and C.L. Williams (2000), "Export-Led Growth: A Survey of the Empirical Literature and Some Non-Causality Results. Part 1 and 2," *J. Int. Trade & Economic Development* 9: 3 261-337

Granger, C.W.J. (1988), "Some Recent Developments in a Concept of Causality," *Journal of Econometrics*, Vol. 39, 199-211.

Greene W. H. (2000), *Econometric Analysis,* *4*th edition, Prentice Hall Inc.

Hayami, Y. (2001) *Development Economics*, 2nd edition,Oxford University Press, Oxford.

Helpman, E. and Krugman, P.R. (1985), *Market Structure and Foreign Trade.* Cambridge, MA: MIT Press.

Hill, C. (1993), *Learning SAS: A computer Handbook for Econometrics,* John Wiley & Sons, Inc.

Johansen, S. (1988), "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control* 12: 231-254

Johansen, S. and K. Juselius (1990), "Maximum likelihood estimation and inference on cointegration with applications to the demand for money", *Oxford Bulletin of Economics and Statistics* 52: 169-211

Keesing, D.B. (1967), "Outward-looking policies and economic development, " *Economic Journal,* 77, 303–20.

Krueger, A.O. (1985), *The experience and lessons of Asia superexporters* In V. Corbo, A.O. Krueger and F. Ossa (eds) *Export-oriented Development Strategies: The Success of Five Newly Industrialized Countries,* London: Westview Press

Krugman, P.R. (1984), *Import protection as export promotion* In H. Kierzkowski (ed.) *Monopolistic Competition in International Trade,* Oxford: Oxford University Press.

Lancaster, K. (1980), "Intra-industry trade under perfect monopolistic competition," Journal of International Economics 10, 151–75.

Lutkepohl, H. (2005), *New Introduction to Multiple Time Series Analysis, Second Edition*, Springer-Verlag

Marin, D. (1992), "Is the export-led growth hypothesis valid for industrialized countries?," *Review of Economics and Statistics*, 74, 678–88

Pack, H. (1988), "Industrialization and trade In H. Chenery and T.N. Srinivasan (eds)," *Handbook of Development Economics, Volume 1*. Amsterdam: Elsevier.

Perron P. (1990), "Tests of Joint Hypotheses for Time Series Regression With a Unit Root," *Advances in Econometrics*, Vol 8, P135-159

Phillips, P. C.B. (1996) "Econometric Model Determination", *Econometrica 64*, 763-812

Phillips, P. C.B., W. Ploberger (1996) "An Asymptotic Theory of Bayesian Inference for Time Series" *Econometrica* 64, 381-413

Phillips, P. C.B. (2002) "New Unit Root Asymptotics in the Presence of Deterministic Trends", *Journal of Econometrics,* 111 (2002) 323-353

Rambaldi, A. N. and H.E.Doran (1996), "Testing for Granger Non-Causality in Cointegrated Systems Made Easy," *working paper, No. 88 - August 1996, University of New England, Austrilia*

Robinson, P.M., F. Iacone (2005), "Cointegration in Fractional Systems with Deterministic Trends", *Journal of Econometrics,* 129 (2005) 263-298

Shan J., Sun F. (1999), "Export-led growth and the US economy: some further testing," *Applied Economics Letters*, vol. 6, no. 3, pp. 169-172(4)

Sims, C.A. (1980), "Macroeconomics and reality," *Econometrica* 48, 1–48

Dufour, J.-M. and Renault, E. (1998), "Short run and long run causality in time series: theory," *Econometrica* 66, 1099–125.

Toda, H.Y. and Yamamoto, T. (1995), "Statistical inference in vector autoregressions with possibly integrated processes," *Journal of Econometrics* 66, 225–50.

Wooldridge J.M. (2002), *Econometric Analysis of Cross Section and Panel Data,* The MIT Press, Cambridge, Massachusetts

Zapata, H.O. and J.M. Gil (1998), "Cointegration and Causality in International AgriculturalEconomics Research," *Journal of Agricultural Economics,* 1356; 1998:1-9

Zapata, H.O. and A.N. Rambaldi (1997), "Monte Carlo Evidence on Cointegration and Causation," *Oxford Bulletin of Economics and Statistics*, 59, 2(1997)