Center for Research in Econometric Analysis of Time Series

CREATES Research Paper 2008-19

Testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form

Anne Péguin-Feissolle, Birgit Strikholm and Timo Teräsvirta

School of Economics and Management
University of Aarhus
Building 1322, DK-8000 Aarhus C
Denmark

Testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form*

Anne Péguin-Feissolle† Birgit Strikholm‡
Anne.Peguin@univmed.fr Birgit.Strikholm@gmail.com
Timo Teräsvirta§¶
tterasvirta@econ.au.dk

March 2008

Abstract

In this paper we propose a general method for testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form. These tests are based on a Taylor expansion of the nonlinear model around a given point in the sample space. We study the performance of our tests by a Monte Carlo experiment and compare these to the most widely used linear test. Our tests appear to be well-sized and have reasonably good power properties.

Key words: Hypothesis testing, causality.
JEL Classification Code: C22, C51

*This research has been supported by Jan Wallander's and Tom Hedelius's Foundation, Grant No. J02-35, and the Danish National Research Foundation. We wish to thank participants of the FPPE workshop (Helsinki, 2003) and seminar participants at the Econometric Institute, Erasmus University Rotterdam for discussion and comments. We also wish to thank Heather Anderson for useful comments but retain the responsibility for any errors and shortcomings in this work. The views expressed in this work are those of the authors' and do not necessarily represent the official views of Bank of Estonia.

†GREQAM, Centre de la Vieille Charité, 2 rue de la Charité, F-13002 Marseille, France

‡Bank of Estonia, Estonia blvd. 13, EE-15095 Tallinn, Estonia

§Department of Economic Statistics and Decision Support, Stockholm School of Economics, P.O. Box 6501, SE-113 83 Stockholm, Sweden

¶CREATES, School of Economics and Management, University of Aarhus, Building 1322, DK-8000 Aarhus C, Denmark

1 Introduction

In a seminal paper, Granger (1969) introduced an operational definition of causality between two variables. In particular, if the variance of the prediction error of the first variable is reduced by including measurements from the second variable, then the second variable is said to have a causal influence on the first variable. This definition has since formed the starting-point for testing the null hypothesis of one variable not causing the other. Note, however, that prediction in the original definition has in practice come to mean in-sample, not necessarily out-of-sample, prediction. The testing has most often been carried out in the linear framework. Lütkepohl (2005) provides a comprehensive overview and an introduction to the testing procedure. For an example of a genuine out-of-sample application, see Ashley, Granger, and Schmalensee (1980).

During the last decade there has been growing interest in generalizing the test to allow for nonlinear relationships between variables. Baek and Brock (1992) suggested a generalization based on the BDS test described in Brock, Dechert, Scheinkman, and LeBaron (1996); Hiemstra and Jones (1994) proposed another version of this test, relaxing the iid assumption. Bell, Kay, and Malley (1996) developed a procedure for causality testing between two univariate time series using non-parametric regression ("generalized" additive models). The above-mentioned tests are all nonparametric and computationally intense. Skalin and Teräsvirta (1999) proposed a parametric test based on the smooth transition regression model and applied that to a set of long Swedish macroeconomic series. That test is easy to compute, but it relies on specific assumptions about the functional form of the causal relationship. Li (2006) suggested a somewhat similar test allowing for threshold effects and augmenting the autoregressive threshold (or smooth transition) model with covariates. Chen, Rangarjan, Feng, and Ding (2004) extended Granger's idea to nonlinear situations by proposing a procedure based on a local linear approximation of the nonlinear function. Apparently, no asymptotic distribution theory is available for inference in this framework, and the results are only descriptive.

In this paper, we suggest two new tests that require little knowledge of the functional relationship between the two variables. The idea is to globally approximate the potential causal relationship between the variables by a Taylor series expansion, which can be seen as a way of linearizing the testing problem. In that sense, noncausality tests based on a single linear regression form a special case in which the Taylor series expansion approximating the actual relationship is of order one. In other words, our framework nests the linear case. Compared to nonparametric procedures, the tests introduced in this paper are very easy to compute. They are also available in large samples where the computational burden of nonparametric techniques becomes prohibitive. Rech, Teräsvirta, and Tschernig (2001) recently applied this idea to nonlinear variable selection.

The paper is organized as follows. Section 2 contains a description of our noncausality tests. Section 3 reports results of a simulation study: both the size and the power of these tests are investigated by Monte Carlo experiments. Section 4 provides a small study based on long Swedish macroeconomic series and Section 5 concludes.

2 Tests of the Granger Noncausality Hypothesis

2.1 Standard linear Granger noncausality test

We begin by recalling the standard way for testing the linear Granger noncausality hypothesis. In that framework, a series x_t is defined not to (linearly) Granger cause another series y_t (x NGC y) if the null hypothesis

$$\mathrm{H}_0 : \beta_1 = \ldots = \beta_q = 0 \tag{1}$$

holds in

$$y_t = \theta_0 + \theta_1 y_{t-1} + \ldots + \theta_p y_{t-p} + \beta_1 x_{t-1} + \ldots + \beta_q x_{t-q} + \varepsilon_t. \tag{2}$$

We make the following assumptions:

A1. $\{\varepsilon_t\}$ *is a sequence of independent, random normal*$(0, \sigma^2)$ *errors.*

A2. $\{y_t\}$ *is stationary and ergodic under (1), that is, the roots of the lag polynomial* $1 - \sum_{j=1}^{q} \theta_j L^j$ *lie outside the unit circle.*

A3. $\{x_t\}$ *is a weakly stationary and ergodic sequence.*

If $\{x_t\}$ is a linear autoregressive-moving average process, then the process is stationary if and only if the roots of the autoregressive lag polynomial lie outside the unit circle. In the nonlinear case, probabilistic properties, such as stationarity and ergodicity, do not seem to be available except in a few special cases, see Stensholt and Tjøstheim (1987), for example.

Assumption A1 is made to allow maximum likelihood-based inference. Robustifying the inference against non-normal errors is possible, however, but is not considered here. Assumptions A2 and A3 guarantee the existence of the second moments needed for the standard distribution theory to be valid.

Under these assumptions one can test the noncausality hypothesis in the single equation framework (2) using an LM statistic (denoted by the subscript SE). Following the recommendation in many earlier papers, we use an F-approximation to the asymptotically χ^2-distributed statistic:

$$Linear_{SE} = \frac{(SSR_0 - SSR_1)/q}{SSR_1/(T-p-q-1)} \overset{approx}{\underset{\mathrm{H}_0}{\sim}} F_{q,T-p-q-1}, \tag{3}$$

where SSR_0 and SSR_1 are sums of squared residuals from regressions under the null and the alternative hypotheses, respectively, and T is the number of observations. The test for testing the null of y_t not Granger causing x_t (y NGC x) can be defined analogously.

Testing the noncausality hypothesis within (2) contains the implicit assumption that y_t does not Granger cause x_t. If this assumption is not valid, then, at least in principle, testing has to be carried out within a bivariate system. Testing the hypothesis of x_t not causing y_t amounts to testing

$$\mathrm{H}_0 : \beta_{11} = \ldots = \beta_{1q_y} = 0 \tag{4}$$

in the system:

$$\begin{aligned} y_t &= \theta_{10} + \theta_{11}y_{t-1} + \ldots + \theta_{1p_y}y_{t-p_y} + \beta_{11}x_{t-1} + \ldots + \beta_{1q_y}x_{t-q_y} + \varepsilon_{yt} \\ x_t &= \theta_{20} + \theta_{21}y_{t-1} + \ldots + \theta_{2p_x}y_{t-p_x} + \beta_{21}x_{t-1} + \ldots + \beta_{2q_x}x_{t-q_x} + \varepsilon_{xt}, \end{aligned} \tag{5}$$

where ε_{it} are assumed white noise with a variance-covariance matrix $\Sigma = \begin{pmatrix} \sigma^2_{yy} & \sigma^2_{yx} \\ \sigma^2_{xy} & \sigma^2_{xx} \end{pmatrix}$ under H_0.

The F-version of the LM-test, see Bewley (1986), for testing (4) within the equation system with feedback (5), denoted by the subscript FB, can then be computed as

$$Linear_{FB} = \frac{T}{q_y}\left(m - \mathrm{tr}\,(\widehat{\mathbf{\Omega}}_{\mathbf{1}}\widetilde{\mathbf{\Omega}}_{\mathbf{0}}^{-1})\right) \overset{approx}{\underset{\mathrm{H}_0}{\sim}} F_{q_y,T}, \tag{6}$$

where m is the number of equations in the system. Matrices $\widetilde{\mathbf{\Omega}}_{\mathbf{0}} = \widetilde{\mathbf{E}}_0{}'\widetilde{\mathbf{E}}_0$ and $\widehat{\mathbf{\Omega}}_{\mathbf{1}} = \widehat{\mathbf{E}}_1{}'\widehat{\mathbf{E}}_1$ are the cross-product matrices of the residuals from estimating the model under the null and under the alternative, respectively. More specifically, $\widetilde{\mathbf{E}}_0 = (\tilde{\boldsymbol{\varepsilon}}_1, \ldots, \tilde{\boldsymbol{\varepsilon}}_T)'$ and $\widehat{\mathbf{E}}_1 = (\hat{\boldsymbol{\varepsilon}}_1, \ldots, \hat{\boldsymbol{\varepsilon}}_T)'$, where $\tilde{\boldsymbol{\varepsilon}}_t$ and $\hat{\boldsymbol{\varepsilon}}_t$, $t = 1, \ldots, T$, are the $(m \times 1)$ residual vectors from the restricted and the unrestricted model, respectively. Analogously, the hypothesis y NGC x corresponds to $\mathrm{H}_0 : \theta_{21} = \ldots = \theta_{2p_x} = 0$ in (5).

2.2 Framework for the general test

Suppose now that we have two weakly stationary and ergodic time series $\{x_t\}$ and $\{y_t\}$. The functional form of the relationship between the two is unknown, but it is assumed that the possible causal relationship between y and x is adequately represented by the following equation:

$$y_t = f_y(y_{t-1}, \ldots, y_{t-p}, x_{t-1}, \ldots, x_{t-q}; \boldsymbol{\theta}) + \varepsilon_{yt}, \tag{7}$$

where $\boldsymbol{\theta}$ is a parameter vector and $\varepsilon_{yt} \sim \mathrm{nid}(0, \sigma_y^2)$. In this framework, x does not Granger cause y if

$$f_y(y_{t-1}, \ldots, y_{t-p}, x_{t-1}, \ldots, x_{t-q}; \boldsymbol{\theta}) = f^*(y_{t-1}, \ldots, y_{t-p}; \boldsymbol{\theta}^*). \tag{8}$$

This means that the conditional mean of y_t given the past values of x_t and y_t is not a function of the past values of x_t.

If the possibility of causality from y to x cannot be excluded *a priori*, one has to assume that there exists a reduced form of the relationship between the two variables. Its precise form is unknown but we assume that it is represented by the following bivariate system:

$$\begin{aligned} y_t &= f_y(y_{t-1}, \ldots, y_{t-p_y}, x_{t-1}, \ldots, x_{t-q_y}; \boldsymbol{\theta}_y) + \varepsilon_{yt} \\ x_t &= f_x(y_{t-1}, \ldots, y_{t-p_x}, x_{t-1}, \ldots, x_{t-q_x}; \boldsymbol{\theta}_x) + \varepsilon_{xt}, \end{aligned} \tag{9}$$

where $\boldsymbol{\theta}_i$, $i = y, x$, are parameter vectors and $\varepsilon_{it} \sim \mathrm{nid}(0, \sigma_i^2)$ and $E\varepsilon_{yt}\varepsilon_{xs} = 0$ for all t, s. In this framework, x NGC y if

$$f_y(y_{t-1}, \ldots, y_{t-p_y}, x_{t-1}, \ldots, x_{t-q_y}; \boldsymbol{\theta}_y) = f^*(y_{t-1}, \ldots, y_{t-p_y}; \boldsymbol{\theta}_y^*) \tag{10}$$

in (9). Analogously, y NGC x if

$$f_x(y_{t-1}, \ldots, y_{t-p_x}, x_{t-1}, \ldots, x_{t-q_x}; \boldsymbol{\theta}_x) = f^{**}(x_{t-1}, \ldots, x_{t-q_x}; \boldsymbol{\theta}_x^*) \tag{11}$$

in (9).

2.3 Noncausality tests based on a Taylor series approximation

The null hypothesis of no Granger causality can be tested as follows. First, linearize f_y and f_x by approximating them with general polynomials. After merging terms and reparameterizing, the kth-order Taylor approximation of f_y has the following form:

$$\begin{aligned} y_t &= \beta_0 + \sum_{j=1}^{p_y} \beta_j y_{t-j} + \sum_{j=1}^{q_y} \gamma_j x_{t-j} + \\ &+ \sum_{j_1=1}^{p_y} \sum_{j_2=j_1}^{p_y} \beta_{j_1 j_2} y_{t-j_1} y_{t-j_2} + \sum_{j_1=1}^{p_y} \sum_{j_2=1}^{q_y} \delta_{j_1 j_2} y_{t-j_1} x_{t-j_2} + \\ &+ \sum_{j_1=1}^{q_y} \sum_{j_2=j_1}^{q_y} \gamma_{j_1 j_2} x_{t-j_1} x_{t-j_2} + \cdots + \\ &+ \sum_{j_1=1}^{p_y} \sum_{j_2=j_1}^{p_y} \cdots \sum_{j_k=j_{k-1}}^{p_y} \beta_{j_1 \ldots j_k} y_{t-j_1} \cdots y_{t-j_k} + \cdots + \\ &+ \sum_{j_1=1}^{q_y} \sum_{j_2=j_1}^{q_y} \cdots \sum_{j_k=j_{k-1}}^{q_y} \gamma_{j_1 \ldots j_k} x_{t-j_1} \cdots x_{t-j_k} + \epsilon_{yt}, \\ &= T_y^k(y,x) + \epsilon_{yt}, \end{aligned} \tag{12}$$

where $\epsilon_{yt} = \varepsilon_{yt} + f_y - T_y^k(y,x)$, and $q_y \leq k$ and $p_y \leq k$ for notational convenience. Expansion (12) contains all possible combinations of lagged values of y_t and lagged values of x_t up to order k. A similar expression can be defined for x_t, and the testing is carried out within the system

$$\begin{cases} y_t = T_y^k(y,x) + \epsilon_{yt} \\ x_t = T_x^k(x,y) + \epsilon_{xt}, \end{cases} \tag{13}$$

where $T_x^k(x,y)$ and ϵ_{xt} are defined analogously.

2.3.1 General test

Owing to the approximation (12), the testing problem is straightforward as it has been returned to the problem of testing a linear hypothesis in a bivariate system that is linear in parameters. The assumption that x_t does not Granger cause y_t means that all terms involving functions of lagged values of x_t in (12) must have zero coefficients. In the most general case[1], the null hypothesis of x_t not Granger causing y_t can be written

[1]We are only going to consider the bivariate case. Extensions to higher-dimensional systems are straightforward.

as

$$
\mathrm{H}_{02}: \begin{cases} \gamma_j = 0, & j = 1, \ldots, q_y \\ \delta_{j_1 j_2} = 0, & j_1 = 1, \ldots, p_y, \quad j_2 = 1, \ldots, q_y \\ \gamma_{j_1 j_2} = 0, & j_1 = 1, \ldots, q_y, \quad j_2 = j_1, \ldots, q_y \\ \vdots \\ \gamma_{j_1 \ldots j_k} = 0, & j_1 = 1, \ldots, q_y, \quad j_2 = j_1, \ldots, q_y, \ \ldots, j_k = j_{k-1}, \ldots, q_y. \end{cases} \tag{14}
$$

We make the following assumptions:

A4. In (9), $\{(\varepsilon_{yt}, \varepsilon_{xt})'\} \sim \mathrm{NID}(0, \Sigma)$.

A5. Sequences $\{x_t\}$ and $\{y_t\}$ are weakly stationary and, in addition, $E(y_t^j x_t^{k-j})^2 = c_k^{(j)} < \infty$, for $j = 0, 1, ..., k$.

A6. $\Pr\{|(1/T)\sum_{t=1}^{T} y_t^{2j} x_t^{2(k-j)} - c_k^{(j)}| > \delta\} < \varepsilon_\delta$ for any $\delta > 0$ and $\varepsilon_\delta > 0$ and for $j = 0, 1, ..., k$, as $T \to \infty$.

A7. $X = \{x_t : x_t \in X\}$, $X \subset \mathbb{R}$, $Y = \{y_t : y_t \in Y\}$, $Y \subset \mathbb{R}$, are compact sets.

A8. Functions $f_z(y_{t-1}, ..., y_{t-p_z}, x_{t-1}, ..., x_{t-q_z})$, $z = x, y$, are continuous and real-valued.

We assume that in the Taylor approximation, the lag lengths p_z, $z = x, y$, are fixed. Furthermore, k, the order of the general polynomial, is independent of T. Then we have the following result.

Theorem 1 *The LM statistic of H_{02} has an asymptotic χ^2 -distribution with the number of degrees of freedom equal to the number of coefficients in H_{02}, when the null hypothesis holds.*

Proof. See the Appendix.

Despite the asymptotic result, this testing problem is in practice a finite-sample problem, which means that determining the lag lengths and the degree of approximation k requires careful attention. In particular, the size of the null hypothesis increases quite rapidly with p_y, q_y and k. For this reason, the F-version of the LM test should be used as it has better size properties in small and moderate samples than the asymptotically valid χ^2-statistic. Thus we apply

$$General_{FB} = \frac{T}{N_1}\left(m - \text{tr}\,(\widehat{\boldsymbol{\Omega}}_{\mathbf{1}}\widetilde{\boldsymbol{\Omega}}_{\mathbf{0}}^{-1})\right) \underset{\text{H}_0}{\overset{approx}{\sim}} F_{N_1,T},$$

where the matrix $\widetilde{\boldsymbol{\Omega}}_{\mathbf{0}} = \widetilde{\mathbf{E}}_0{}'\widetilde{\mathbf{E}}_0$ is obtained from regression (13) under the null hypothesis and $\widehat{\boldsymbol{\Omega}}_{\mathbf{1}} = \widehat{\mathbf{E}}_1{}'\widehat{\mathbf{E}}_1$ from the full regression (13). Furthermore, T is the number of observations and N_1 the number of parameters in (12) to be tested under the null hypothesis. The latter quantity is defined as follows:

$$N_1 = N - N_2 = \left(1 + \sum_{r=1}^{k}\binom{p_y + q_y + r - 1}{r}\right) - \left(1 + \sum_{r=1}^{k}\binom{p_y + r - 1}{r}\right), \qquad (15)$$

where N is the total number of parameters and N_2 the number of parameters not under test.

There are two practical difficulties related to equation (13). One is numerical whereas the other one has to do with the amount of information. Numerical problems may arise because the regressors in (13) tend to be highly collinear if k, p_y and q_y (and also p_x, q_x) are large. The other difficulty is, as already mentioned, that the number of regressors increases rapidly with k, so the dimension of the null hypothesis may become rather large. For instance, when $p_y = 2$ and $q_y = 3$ then $N_1 = 46$ for $k = 3$, and $N_1 = 231$ when $k = 5$. A practical solution to both problems is to replace some matrices by their largest principal components as follows. First, divide the regressors in the auxiliary test-equation (12), say, into two groups: those being functions of lags of y_t only and the remaining ones. Replace the second group of regressors by their first p^* principal components. The null hypothesis now is that the principal components have zero coefficients. This yields the following test statistic:

$$General^*_{FB} = \frac{T}{p^*}\left(m - \text{tr}\,(\widehat{\boldsymbol{\Omega}}_{\mathbf{1}}\widetilde{\boldsymbol{\Omega}}_{\mathbf{0}}^{-1})\right) \underset{\text{H}_0}{\overset{approx}{\sim}} F_{p^*,T}. \qquad (16)$$

The 'remainder' now also includes the approximation error due to the omitted principal components related to the smallest eigenvalues.

2.3.2 Semi-additive test

In some cases it may be reasonable to assume that the general model is "semi-additive". That means that the model has the following form:

$$\begin{aligned} y_t &= g_y(y_{t-1},\dots,y_{t-p_y};\boldsymbol{\theta}_{yy}) + f_y(x_{t-1},\dots,x_{t-q_y};\boldsymbol{\theta}_{yx}) + \varepsilon_{yt} \\ x_t &= g_x(y_{t-1},\dots,y_{t-p_x};\boldsymbol{\theta}_{xy}) + f_x(x_{t-1},\dots,x_{t-q_x};\boldsymbol{\theta}_{xx}) + \varepsilon_{xt}. \end{aligned} \tag{17}$$

Here we assume that f_y, f_x, g_y and g_x satisfy the assumptions similar to the ones for f_y and f_x in A8. Assumption A5 can now weakened as follows:

A5′. $Ey_t^{2k} = c_{yk} < \infty, \qquad Ex_t^{2k} = c_{xk} < \infty.$

We state again that x_t does not cause y_t if the past values of x_t contain no information about y_t that is not already contained in the past values of y_t. When this is the case $f_y(x_{t-1},\dots,x_{t-q_y};\boldsymbol{\theta}_{yx}) \equiv \textit{constant}$. The functions g_y, g_x, f_y and f_x are now separately expanded into kth-order Taylor series. For example, linearizing g_y and f_y in (17) by expanding both functions into a kth-order Taylor series around arbitrary points in the sample space, merging terms and reparameterizing, yields

$$\begin{aligned} y_t &= \beta_0 + \sum_{j=1}^{p_y}\beta_j y_{t-j} + \sum_{j=1}^{q_y}\gamma_j x_{t-j} + \sum_{j_1=1}^{p_y}\sum_{j_2=j_1}^{p_y}\beta_{j_1j_2}y_{t-j_1}y_{t-j_2} + \\ &+ \sum_{j_1=1}^{q_y}\sum_{j_2=j_1}^{q_y}\gamma_{j_1j_2}x_{t-j_1}x_{t-j_2} + ... + \sum_{j_1=1}^{p_y}\sum_{j_2=j_1}^{p_y}...\sum_{j_k=j_{k-1}}^{p_y}\beta_{j_1...j_k}y_{t-j_1}...y_{t-j_k} + \\ &+ \sum_{j_1=1}^{q_y}\sum_{j_2=j_1}^{q_y}...\sum_{j_k=j_{k-1}}^{q_y}\gamma_{j_1...j_k}x_{t-j_1}...x_{t-j_k} + \epsilon_t, \end{aligned} \tag{18}$$

where $q_y \le k$ and $p_y \le k$ for notational convenience. Expansion (18) contains all possible combinations of y_{t-j} and x_{t-i} up to order k, but no cross-terms. Therefore, the hypothesis x NGC y becomes:

$$H_{03}: \left\{ \begin{aligned} \gamma_j &= 0,\ j=1,...,q_y \\ \gamma_{j_1j_2} &= 0,\ j_1=1,...,q_y,\ j_2=j_1,...,q_y \\ &\vdots \\ \gamma_{j_1...j_k} &= 0,\ j_1=1,...,q_y,\ j_2=j_1,...,q_y,...,\ j_k=j_{k-1},...,q_y. \end{aligned} \right.$$

The number of parameters to be tested under the null hypothesis is

$$N_{11} = \sum_{r=1}^{k}\binom{q_y+r-1}{r}.$$

Although the number of parameters for any fixed k is smaller than in the unrestricted nonadditive case, the problems of collinearity and the large dimension of the null hypothesis may still be present. The previous solution still applies: the regressors are replaced by p^* principal components of corresponding observation matrix. Again an LM-type test can be used, and the resulting test statistic is called $Additive^*_{FB}$. Under the null hypothesis, $Additive^*_{FB}$ has approximately an F-distribution with p^* and T degrees of freedom. Note that approximating f_y through principal components may only affect the power of the test, not its size.

If equation (17) is valid, then the corresponding test can be expected to be more powerful than the ones based on equation (9). Conversely, applying $Additive^*_{FB}$ when the relationship is not semi-additive as in (17) may result in a substantial loss of power compared to the power of $General^*_{FB}$.

2.4 Global vs. local approximation of the nonlinear system

As discussed in an earlier section, our approach is based on global approximation of the unknown nonlinear function. The starting-point for the local linear approximation of Chen et al. (2004) is the standard delay coordinate embedding reconstruction of the phase space attractors, see Boccaletti, Valladares, Pecora, Geffert, and Carroll (2002). A full description of a given attractor requires a nonlinear set of equations, but it is possible to locally approximate the dynamics linearly by a vector autoregressive model. Chen et al. then test for Granger causality at each local neighbourhood, average the resulting statistical quantities over the entire attractor and compute the so-called Extended Granger Causality Index. A number of decisions have to be made when using their method: one has to determine the embedding dimension and time delay. Determining the optimal neighbourhood size is also a nontrivial issue. It appears that no asymptotic distribution theory is available for inference in this framework, so the results are obviously bound to be rather descriptive. It may be guessed that an application of this procedure requires rather long time series unless the nonlinear relationship is nearly linear.

3 Monte Carlo Experiments

3.1 Simulation design

In this section we shall investigate the small-sample performance of the proposed non-causality tests. We compare the tests with the standard linear testing procedure because that is what practitioners generally use in their work. Moreover, it is often the case that the researcher chooses to ignore the possible presence of feedback (causality in the other direction) and conducts the analysis within a single equation. One may then ask: does it matter whether the presence of feedback is explicitly acknowledged or not? On the one hand this should matter, and consequently the tests should be carried out in a system framework using (13) or in an equivalent semi-additive system. On the other hand, the restrictions implied on the system by the null hypothesis are not cross-equation restrictions, which suggests that testing could be carried out by only using one equation of the system.

We report the results for all tests based both on the bivariate equation system (denoted with subscript FB) and on the single equation only (denoted with subscript SE), and compare the results. The tests included in our comparison are:

- $Linear_{SE}$ and $Linear_{FB}$ defined in section (2.1), formulas (3), (6)
- $General^{*}_{SE}$ and $General^{*}_{FB}$ defined in section (2.3.1), formula (16)
- $Additive^{*}_{SE}$ and $Additive^{*}_{FB}$ defined in section (2.3.2).

We present our size and power results for all tests graphically as Davidson and MacKinnon (1998) have recommended. Their graphs are easier to interpret than the conventional tables typically used for reporting such results. The basis of these graphs is the empirical distribution function (EDF) of the p-values of the simulated realizations τ_j, $j = 1, ..., N$, of some test statistic τ. Let p_j be the p-value associated to τ_j, i.e., $p_j \equiv p(\tau_j) = P(\tau > \tau_j)$, the probability of observing a value of τ greater than τ_j of the statistic. The EDF of the p_j's is defined by:

$$\widehat{F}(\xi_i) = \frac{1}{N}\sum_{j=1}^{N} I(p_j \leq \xi_i), \tag{19}$$

where I is an indicator function given by :

$$I(p_j \leq \xi_i) = \begin{cases} 1 & \text{if } p_j \leq \xi_i \\ 0 & \text{otherwise} \end{cases}$$

and ξ_i is a point in the $[0,1]$ interval. Following Davidson and MacKinnon (1998), a parsimonious set of values ξ_i, $i = 1, ..., m$, is

$$\xi_i = 0.002, 0.004, ..., 0.01, 0.02, ..., 0.99, 0.992, ..., 0.998 \quad (m = 107). \tag{20}$$

Concerning the size of the tests, we present the simple plot of $\widehat{F}(\xi_i) - \xi_i$ against ξ_i for each test. We know that if the distribution of τ is the one assumed under the null hypothesis, the p_j's should be a sample from a uniform $[0,1]$ distribution. In that case, the plot of $\widehat{F}(\xi_i)$ against ξ_i should be close to the 45° line, whereas $\widehat{F}(\xi_i) - \xi_i$ should fluctuate around zero as a function of ξ_i.

The results of the power comparisons are reported using simple power curves, instead of the size-corrected size-power curves advocated by Davidson and MacKinnon. The reason for this is that in practice we would not know how to size-correct the results, and our aim is to study the tests from the practitioners point of view. Therefore, we graph the locii of points $((\xi_i), \widehat{F}^*(\xi_i))$ where the values of the ξ_i's are given by (20), and $\widehat{F}^*(\xi_i)$ is the EDF generated by a process belonging to the alternative hypothesis. In other words, we record the p-values for every Monte Carlo replication and just plot the curves corresponding to rejection rates at given nominal levels.

3.2 Simulation results

For all the experiments, the number of replications $N_R = 1000$ and the number of observations[2] $T = 150$. The innovations $\varepsilon_{it} \sim \text{nid}(0,1)$, $i = y, x$, $t = 1, \ldots, T$, and sequences $\{\varepsilon_{yt}\}$ and $\{\varepsilon_{xt}\}$ are mutually independent in all experiments. We make use of the second-order Taylor approximation of f_y, g_y, f_x and g_x where $p_y = 3$; $q_y = 3$; $q_2 = 3$; $p_x = 3$. The number of principal components is determined separately for each case. Only the largest principal components that together explain at least 90% of the variation in the matrix of observations are used. The system consists of unrelated equations and is estimated equation by equation by least squares.

For every DGP we present two graphs: panel (a) contains the results of the test of x NGC y, and panel (b) the results of the test of y NGC x. In every panel the performance of the three causality tests *Linear*, *General** and *Additive**, is reported, both for the single equation (SE) and the system (FB) framework. The corresponding lines on

[2]We let the data-generating process run for a while to eliminate the possible initial effects, i.e., we discard first 500 observations and use only the last 150.

graphs are labelled Linear_SE, General*_SE, Additive*_SE, Linear_FB, General*_FB and Additive*_FB, respectively.

Empirical size of the tests

To illustrate the behaviour of the tests under the null hypothesis, we simulated six different systems. The data-generating processes are presented together with the p-values of the linearity test of Harvill and Ray (1999), denoted by HRp. These p-values are reported in order to give the reader an indication of how nonlinear the systems are. A very small p-value indicates that the system is strongly nonlinear, whereas larger values suggest only weak or no nonlinearity.

The first system is a first-order vector autoregressive model ($HRp = 0.59$):

$$\begin{aligned} y_t &= 1 + 0.3y_{t-1} + 0.1\varepsilon_{yt} \\ x_t &= 0.4 - 0.63x_{t-1} + 0.2\varepsilon_{xt}. \end{aligned} \tag{21}$$

The second experiment involves a nonlinear system where y_t is generated by a logistic smooth transition autoregressive (STAR) model and x_t follows a bilinear model ($HRp = 8 \times 10^{-10}$):

$$\begin{aligned} y_t &= (0.02 - 0.9y_{t-1} + 0.795y_{t-2})/(1 + \exp(-25(y_{t-1} - 0.02))) + 0.1\varepsilon_{yt} \\ x_t &= 0.8 - 0.6x_{t-1} + 0.1\varepsilon_{x,t-1}x_{t-1} + 0.3\varepsilon_{xt}. \end{aligned} \tag{22}$$

In the third system y_t is ratio-polynomial and x_t is generated by an exponential STAR model ($HRp = 0.054$):

$$\begin{aligned} y_t &= -0.8 + 0.9/(1 + y_{t-1}^2 + y_{t-2}^2) + 0.1\varepsilon_{yt} \\ x_t &= (0.2 - 0.6x_{t-1} + 0.45x_{t-2})(1 - \exp(-10(x_{t-1})^2)) + 0.1\varepsilon_{xt}. \end{aligned} \tag{23}$$

The fourth system consists of two self-exciting threshold autoregressive (SETAR) models. Note that this model is not covered by Theorem 1, because the SETAR model does not satisfy Assumption A8. Nevertheless, it is interesting to see how the test behaves in this situation. The system has the following form ($HRp = 2 \times 10^{-30}$):

$$\begin{aligned} y_t &= \begin{cases} 0.1y_{t-1} + \varepsilon_{yt} & y_{t-1} \le 0 \\ -0.5y_{t-1} + \varepsilon_{yt} & y_{t-1} > 0 \end{cases} \\ x_t &= \begin{cases} -0.5 + 0.5x_{t-1} - 0.7x_{t-2} + \varepsilon_{xt} & x_{t-1} \le 0 \\ 0.5 - 0.3x_{t-1} + 0.2x_{t-2} + \varepsilon_{xt} & x_{t-1} > 0\,. \end{cases} \end{aligned} \tag{24}$$

The fifth system is linear with fifth-order autoregression such that causality is only running in one direction, from y to x ($HRp = 0.55$):

$$\begin{aligned} y_t &= 1.41 y_{t-1} - 1.38 y_{t-2} + 1.0813 y_{t-3} - 0.23015 y_{t-4} + 0.0182 y_{t-5} + \varepsilon_{yt} \\ x_t &= 1 - 0.55 x_{t-1} + 0.16 x_{t-2} - 0.4 y_{t-4} - 0.3 y_{t-5} + \varepsilon_{xt}. \end{aligned} \quad (25)$$

The final system is a bivariate nonlinear MA model ($HRp = 2 \times 10^{-12}$):

$$\begin{aligned} y_t &= \varepsilon_{yt} - 0.4\varepsilon_{y,t-1} + 0.3\varepsilon_{y,t-2} + 0.4\varepsilon^2_{y,t-1} \\ x_t &= \varepsilon_{xt} + 0.55\varepsilon_{x,t-1} - 0.3\varepsilon_{x,t-2} - 0.2\varepsilon^2_{x,t-1}. \end{aligned} \quad (26)$$

The results appear in Figures 1 – 6. They show the p-value discrepancy plots, i.e. the graphs of $\widehat{F}(\xi_i) - \xi_i$ against ξ_i. These figures are reproduced for small nominal sizes that are of practical interest.

The size distortions seem generally minor at low levels of significance. The single equation-based tests seem somewhat less size distorted than the system-based ones. Also the *Linear* test seems better-sized than the *General* or *Additive* tests, unless there is feedback, as in Equation (25) is which case the *Linear* test is grossly oversized. Size distortions seen in Figure 5 occur partly because of the misspecified lag length under the null hypothesis: only three lags are used in Taylor expansions. But, there is feedback from y to x through fourth and fifth lag, and when lags of x enter the auxiliary test-equation, they are found to be helpful in explaining y. The same explanation - too few lags used in the approximation - is valid when explaining the size distortions for the nonlinear moving average model. These size distortions can partly be removed by using more lags when approximating the possibly nonlinear causal relationship. The size distortions of test statistics *General* and *Additive* are much smaller for simple linear systems when the order of Taylor expansions is low, i.e. the approximation nests the DGP and there are fewer nuisance auxiliary terms in the test equations. Naturally, the size distortions diminish when the error variance is reduced and when the true lag length is used in the Taylor approximation. We recommend first testing the linearity of the system as in Harvill and Ray (1999), and if linearity is not rejected, using the *Linear* test should suffice.

We conducted additional experiments with the six DGPs above by letting the error covariance matrix differ from the identity matrix. This represents a situation where the original assumptions are violated and there exists a third common factor that simultaneously affects x and y. We let $\Sigma = \begin{pmatrix} \sigma_{yy}^2 & \sigma_{yx}^2 \\ \sigma_{xy}^2 & \sigma_{xx}^2 \end{pmatrix}$ where $\sigma_{xy} \neq 0$. Instead of specifying the exact covariance structure we set the correlation between respective variables to be $\rho = \{-0.9, -0.6, -0.3, 0.3, 0.6, 0.9\}$. For the first three DGPs (the error terms in these equations are multiplied by small coefficients) the size distortions remain about the same or increase slightly for large correlations. For the remaining systems (the effect of error terms is not reduced by a factor) the increase in the size distortion is huge, particularly when the correlation between the error terms is positive and large.

Empirical power of the tests

In this subsection we consider a number of cases where (nonlinear) Granger causality is present between the variables. More precisely, the power-curve figures correspond to the following systems:

- Figure 7 ($x \rightarrow y$ bilinear; $y \rightarrow x$ linear, $HRp = 7 \times 10^{-7}$):

$$\begin{aligned} y_t &= 0.5 + 0.1 y_{t-1} + 0.5 \varepsilon_{y,t-1} x_{t-1} + \varepsilon_{yt} \\ x_t &= 0.22 - 0.39 x_{t-1} + 0.46 y_{t-2} + \varepsilon_{xt} \end{aligned} \tag{27}$$

- Figure 8 ($x \rightarrow y$ (long)linear[3]; $y \rightarrow x$ linear, $HRp = 0.847$):

$$\begin{aligned} y_t &= 1.1 y_{t-1} - 0.56 y_{t-2} + 0.1591 y_{t-3} - 0.0119 y_{t-4} + 0.55 x_{t-4} + 0.1 \varepsilon_{yt} \\ x_t &= 0.5 - 0.1566 x_{t-1} + 0.2083 x_{t-2} - 0.4 y_{t-2} + 0.3 \varepsilon_{xt} \end{aligned} \tag{28}$$

- Figure 9 ($y \rightarrow x$ (long)linear, $HRp = 0.575$):

$$\begin{aligned} y_t &= 1.41 y_{t-1} - 1.38 y_{t-2} + 1.08 y_{t-3} - 0.23 y_{t-4} + 0.02 y_{t-5} + 0.5 \varepsilon_{yt} \\ x_t &= 1 - 0.55 x_{t-1} + 0.16 x_{t-2} - 0.4 y_{t-4} - 0.3 y_{t-5} + 0.5 \varepsilon_{xt} \end{aligned} \tag{29}$$

- Figure 10 ($x \rightarrow y$ semi-additive, $HRp = 0.243$):

$$\begin{aligned} y_t &= 0.1 + 0.4 y_{t-2} + (0.5 - 0.8 x_{t-1})/(1 + \exp(-5(x_{t-3} - 0.2))) + 0.5 \varepsilon_{yt} \\ x_t &= 0.22 + 0.39 x_{t-1} - 0.55 x_{t-2} + 0.3 \varepsilon_{xt} \end{aligned} \tag{30}$$

[3] By "(long)" we denote the situation where the lag length in the actual DGP is longer than the one used in approximations.

- Figure 11 ($x \to y$ semi-additive; $y \to x$ (long)linear, $HRp = 6 \times 10^{-30}$):

$$\begin{aligned} y_t &= 0.5y_{t-1} - 0.6x_{t-1}^2/(1 + x_{t-2}^2) + \varepsilon_{yt} \\ x_t &= 1 + 0.4x_{t-1} - 0.3x_{t-2} + 0.4y_{t-4} - 0.3y_{t-5} + 0.2y_{t-6} + \varepsilon_{xt} \end{aligned} \tag{31}$$

- Figure 12 ($x \to y$ general, $HRp = 8 \times 10^{-23}$):

$$\begin{aligned} y_t &= \begin{cases} -1 + 0.5y_{t-1} - 0.9y_{t-2} + \varepsilon_{yt} & x_{t-1} \leq -0.2 \\ \;\;2 + 0.3y_{t-1} + 0.2y_{t-2} + \varepsilon_{yt} & x_{t-1} > -0.2 \end{cases} \\ x_t &= 0.2 - 0.56x_{t-1} + \varepsilon_{xt} \end{aligned} \tag{32}$$

- Figure 13 ($x \to y$ general, $HRp = 0.00883$):

$$\begin{aligned} y_t &= 0.1 + 0.4y_{t-2} + (1 - 0.8y_{t-2})/(1 + \exp(-9x_{t-1}^2)) + 0.15\varepsilon_{yt} \\ x_t &= 0.22 + 0.39x_{t-1} - 0.55x_{t-2} + 0.3\varepsilon_{xt} \end{aligned} \tag{33}$$

- Figure 14 ($x \to y$ general; $y \to x$ linear, $HRp = 1 \times 10^{-5}$):

$$\begin{aligned} y_t &= 1 - 0.2y_{t-1} + (-1 + 0.5y_{t-2})(1 - \exp(-10x_{t-1}^2)) + 0.3\varepsilon_{yt} \\ x_t &= 0.5x_{t-1} + 0.3y_{t-1} + 0.5\varepsilon_{xt} \end{aligned} \tag{34}$$

- Figure 15 ($x \to y$ bilinear; $y \to x$ general, $HRp = 9 \times 10^{-24}$):

$$\begin{aligned} y_t &= 0.1y_{t-1} - 0.5y_{t-2} + 0.4x_{t-1}\varepsilon_{y,t-1} + 0.5x_{t-2}\varepsilon_{y,t-2} + 0.3\varepsilon_{yt} \\ x_t &= 1 + 0.3x_{t-1} - 0.5x_{t-2} + (2 + 0.4x_{t-1} - 0.3x_{t-2} - 0.15x_{t-3}) \\ &\quad (1 - \exp(-10y_{t-2}^2)) + 0.1\varepsilon_{xt} \end{aligned} \tag{35}$$

- Figure 16 ($x \to y$ general; $y \to x$ semi-additive, $HRp = 8 \times 10^{-5}$):

$$y_t = \begin{cases} 0.1y_{t-1} + 0.9x_{t-1}^2 + 0.4\varepsilon_{yt} & y_{t-1} \leq 0 \\ -0.5y_{t-1} + 0.4\varepsilon_{yt} & y_{t-1} > 0 \end{cases} \tag{36}$$

$$x_t = 0.3x_{t-1} + \frac{0.9}{(1 + x_{t-1}^2 + x_{t-2}^2)} - \frac{0.5}{(1 + \exp(-2y_{t-1}))} + 0.25\varepsilon_{xt}$$

- Figure 17 ($x \to y$ general; $y \to x$ general, $HRp = 8 \times 10^{-8}$):

$$y_t = \begin{cases} 0.1y_{t-1} + 0.3x_{t-1}^2 - 0.5x_{t-2}^2 + 0.2\varepsilon_{yt} & y_{t-1} \leq 0 \\ -0.3y_{t-1} - 0.5x_{t-1}^2 + 0.7x_{t-2}^2 + 0.2\varepsilon_{yt} & y_{t-1} > 0 \end{cases} \tag{37}$$

$$x_t = 0.3x_{t-1} + \frac{0.9}{(1 + x_{t-1}^2 + x_{t-2}^2)} + \frac{(-0.5x_{t-1} + 0.3x_{t-2})}{(1 + \exp(-30y_{t-1}))} + 0.24\varepsilon_{xt}$$

- Figure 18 ($x \to y$ general; $y \to x$ general, $HRp = 3 \times 10^{-28}$):

$$y_t = \begin{cases} 0.1y_{t-1} + \dfrac{(2 - 0.45x_{t-1})}{(1 + \exp(-5x_{t-1}))} + 0.4\varepsilon_{yt} & y_{t-1} \leq 0 \\ -0.5y_{t-1} + \dfrac{(1 - 0.3x_{t-1} + 0.45x_{t-2})}{(1 + \exp(-5x_{t-1}))} + 0.4\varepsilon_{yt} & y_{t-1} > 0 \end{cases} \quad (38)$$

$$x_t = 1 + 0.3x_{t-1}y_{t-2} - 0.15x_{t-2}y_{t-3} + \varepsilon_{xt}.$$

The results of the *Linear* causality test do not offer any great surprises. It is clear that the test works best when the true causal relationship is linear (Figures 7(b), 8(b), 14(b)) but it may only have weak power when this is no longer the case. The *Additive* test as well as the *General* one both suffer somewhat from overparameterization when applied to linear systems. The *Linear* test also seems to perform well when it comes to detecting slowly evolving logistic STAR-type causal relationships, see Figures 10(a), 16(b). It also works surprisingly well for a case when the causing variable is the threshold variable in a two-regime TAR model, 12(a). Note that in this particular example there seems to be no size distortion in testing y NGC x, although the TAR model does not satisfy Assumption A8. *Linear* test also seems to be able to detect a (linear) causal relationship when the lags contributing to the explanation of the other variable are outside the range of the lags included in Taylor expansion (and thus used in the test), see Figure 11(b), 9(b).

At small nominal sizes, the *Additive* test is the best performer of these three tests in Figure 16(a), where the corresponding model actually is semi-additive. It is often more powerful than the *General* test even when the true model is no longer semi-additive, see Figures 13(a), 14(a), for example. Figures 14(a) and 15(b) illustrate the behaviour of the tests in the case where the causality is represented by an exponential smooth transition regression function and the causing variable is the transition variable. The nonlinearity in those models is of *General*-type, but the semi-additive approximation seems to capture most of the relationship. Consequently the *Additive* test is the most powerful one. From the low power of the *Linear* test it can be inferred that in this case excluding the higher order terms from the auxiliary regression is not a good idea.

Figures 7(a) and 15(a) correspond to systems with a bilinear equation and in those cases the *General* test strongly dominates the other test procedures. This may be expected as the relationship is no longer semi-additive, and making that assumption implies a loss of power. From Figures 17(a) and 18(a), it is seen that the *General* test also seems to perform well in a case where the causing variable enters through a regime or the regimes of a SETAR model.

When interpreting the results, one should be aware of the fact that the power of the tests depends on the variance of the error term $c\varepsilon_t$ which controls the signal-to-noise ratio. Even some of the performance rankings indicated above may be changed by varying c. Also, for the ESTR-type models, the ability of the tests to detect causality depends quite heavily on the presence of the intercept in the nonlinear part. When there is only change in the amplitude of the fluctuations and no clear shift in the mean, the power of all tests is extremely low.

There seems to be no big difference in performance between the single equation and system-based tests. This indicates that not much power is lost by ignoring the possible feedback. This is obviously due to the fact that the restrictions imposed by the null hypothesis are not cross-equation restrictions, so little is gained by including the unrestricted equation in the considerations.

4 Application

In this section we analyse the same data as Skalin and Teräsvirta (1999), that is, nine long annual Swedish macroeconomic time series: Gross Domestic Product (GDP), Industrial Production (IP), Private Consumption (CONS), Investment (INV), Exports (EXP), Imports (IMP), Employment[4] (EMPL), Real Wages (RW) and Productivity[5] (PROD). For most series the data span the period from 1861 to 1988, the productivity and employment series begin in 1870. To guarantee stationarity we work with log-differenced data. The autoregressive order of each model is selected using the Akaike information criterion (AIC) and the Godfrey-Breusch (GB) test of no error autocorrelation. For both methods the maximum lag length is set to twelve. If AIC selects a model with less than three lags but GB points to a model with more than twelve lags (most probably picking up on spurious correlation) we make a compromise and use four autoregressive lags in our model. The selected AR orders are given under the variable names in Table 1 - Table 3, that present the results of the *Linear*, *Additive* and *General* test respectively. The pairwise testing is conducted in the single equation framework.

[4]Measured in worked hours in manufacturing and mining.
[5]Industrial production divided by hours worked.

In the tables the column labelled q gives the lag order for the causing variable. The entries in the tables denote the strength of the rejection of the null hypothesis of no Granger causality: a single $*$ denotes a rejection at 5% level, a double $**$ denotes a rejection at 1% level, $***$ and $****$ denote rejections at 0.1% and 0.01% level, respectively. A rejection of the null hypothesis does not imply a direct causal link between the pair of variables, though. Changing the information set underlying our bivariate tests may change the results. Nevertheless, the tests are suggestive about the extent of interactions between the variables.

Table 1 contains the results from the *Linear* tests of Granger non-causality. Exports seems to be more of a causing than caused variable: it helps predicting six other variables. Each variable in the set is Granger-causing two to three other variables, except for Investments. Real Wages and Employment are most often influenced by other variables. Strongest links are running from Imports to Real Wages and from Real Wages to Employment.

Tables 2 and 3 contain the results for the *Additive* and *General* test, respectively. Here we choose to make use of the second-order Taylor approximation and principal components to be able to estimate the auxiliary model. The third-order approximation turns out infeasible, because with $\hat{p} = 9$ we would have 220 parameters to estimate under the null, but we only have less than 130 observations available. Alternatively, one could use the "economy version" of the test advocated in Luukkonen et al. (1988), i.e. use the third-order expansion and discard some intermediate higher order cross-terms from the auxiliary regression.

The two tests give rather similar results. Largest differences appear for Imports, where the *General* test finds Industrial Production, Gross Domestic Product and Productivity to be useful predictors at more lags and/or with stronger rejections compared to the *Additive* test. Compared to *Linear* test some rejections appear, some disappear. It seems that the added flexibility in explaining a variable through its own past in a nonlinear manner reduces the importance of other variables as predictors. For example, GDP and Consumption lose their importance when it comes to predicting Employment. In a simple linear framework GDP and Consumption are found to Granger-cause Employment, but not (to the same extent) when nonlinear approximation is employed. The opposite happens for Industrial Production. In the linear framework no other variable seems to Granger-cause IP, the nonlinear testing framework is able to identify links to four variables: Productivity, Real Wages, Investments and Employment.

The main conclusion we can draw from these results is the same as Skalin and Teräsvirta (1999) did: the functional form of the model (linear, STAR, general nonlinear) strongly affects the outcome of these tests.

5 Conclusions

The noncausality tests introduced in this paper are based on standard statistical distribution theory. The size simulations indicate that the idea of polynomial approximation of unknown nonlinear functions is applicable in small samples. The right balance between the number of lags, the order of the Taylor expansion, the degree of nonlinearity and the number of observations is important, however. The power simulations suggest that the tests are indeed useful in discovering potential Granger causality between variables. They also demonstrate the obvious fact that the more we know about the functional form, the more we gain in terms of power. If the true causal relationship is nonlinear whereas testing is carried out under the assumption of linearity, the ensuing loss of power may be substantial. It is therefore advisable to test the Granger noncausality hypothesis both in the linear and the nonlinear framework to ensure that existing causal relationships between the variables are found. Because our tests are based on the idea of linearizing the unknown relationship between the variables, they are not computationally more difficult to carry out than traditional linear tests. However, the length of the time series may restrict the applicability of our technique. Given a sufficient amount of data, our tests should be a useful addition to the toolbox of both applied economists and time series econometricians interested in empirical investigations of Granger causality.

References

Ashley, R., C. W. J. Granger, and R. Schmalensee (1980). Advertising and aggregate consumption: An analysis of causality. *Econometrica 48*, 1149–1168.

Baek, E. G. and W. A. Brock (1992). A general test for nonlinear Granger causality: bivariate model. Working paper, Iowa State University and University of Wisconsin, Madison.

Bell, D., J. Kay, and J. Malley (1996). A non-parametric approach to non-linear causality testing. *Economics Letters 51*, 7–18.

Bewley, R. (1986). *Allocation Models*. Cambridge, Massachusetts: Ballinger Publishing Company.

Boccaletti, S., D. L. Valladares, L. M. Pecora, H. P. Geffert, and T. Carroll (2002). Reconstructing embedding spaces of coupled dynamical systems from multivariate data. *Physical Review E 65*, 035204.

Brock, W. A., W. D. Dechert, J. A. Scheinkman, and B. D. LeBaron (1996). A test for independence based on the correlation dimension. *Econometric Reviews 15*, 197–235.

Chen, Y., G. Rangarjan, J. Feng, and M. Ding (2004). Analyzing multiple nonlinear time series with extended Granger causality. *Physics Letters A 324*, 26–35.

Davidson, R. and J. G. MacKinnon (1998). Graphical methods for investigating the size and power of hypothesis tests. *The Manchester School 66*, 1–26.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica 37*, 424–438.

Harvill, J. L. and B. K. Ray (1999). A note on tests for nonlinearity in a vector time series. *Biometrika 86*, 728–734.

Hiemstra, C. and J. D. Jones (1994). Testing for linear and nonlinear Granger causality in stock price-volume relation. *Journal of Finance 49*, 1639–1664.

Li, J. (2006). Testing Granger Causality in the presence of threshold effects. *International Journal of Forecasting 22*, 771–780.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer Verlag, Berlin.

Luukkonen, R., P. Saikkonen, and T. Teräsvirta (1988). Testing linearity against smooth transition autoregressive models. *Biometrika 75*, 491–499.

Rech, G., T. Teräsvirta, and R. Tschernig (2001). A simple variable selection technique for nonlinear models. *Communications in Statistics, Theory and Methods 30*, 1227–1241.

Royden, H. L. (1963). *Real Analysis.* Macmillan, New York.

Skalin, J. and T. Teräsvirta (1999). Another look at Swedish business cycles, 1861-1988. *Journal of Applied Econometrics 14*, 359–378.

Stensholt, B. K. and D. Tjøstheim (1987). Multiple bilinear time series models. *Journal of Time Series Analysis 8*, 221–233.

Appendix

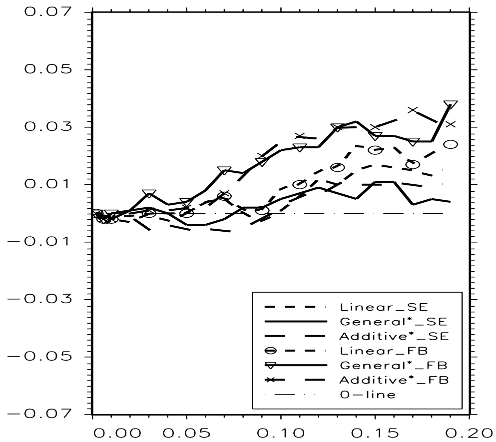
Theorem 1 *The LM statistic of H_{02} has an asymptotic χ^2 -distribution with the number of degrees of freedom equal to the number of coefficients in H_{02}, when the null hypothesis holds.*

Proof. From Assumption A7 it follows that $X^{(q)} = X \times ... \times X \subset \mathbb{R}^q$, $q = \max(q_x, q_y)$, and $Y^{(p)} = Y \times ... \times Y \subset \mathbb{R}^p$, $p = \max(p_x, p_y)$, are compact sets. Then, by Tychonoff's theorem, $X^{(q)} \times Y^{(p)}$ is a compact set. Since by Assumption A8, $f_z(y_{t-1}, ..., y_{t-p_z}, x_{t-1}, ..., x_{t-q_z})$, $z = x, y$, are continuous and real-valued functions, they are also bounded. The same is true for $f_y(y_{t-1}, ..., y_{t-p_y})$.

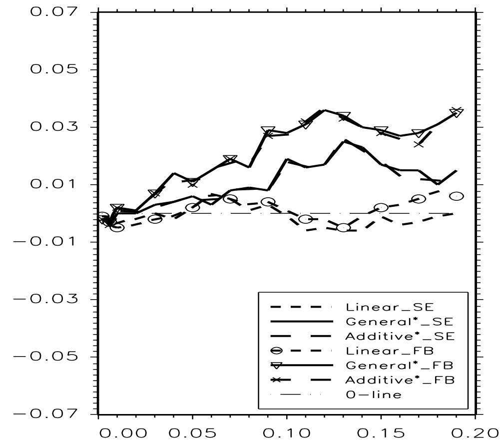
It then follows from a corollary to the Stone-Weierstrass theorem (see Royden (1963), p. 151) that

$$|f_y(y_{t-1}, ..., y_{t-p_y}) - T_y^k(0, y)| < \varepsilon$$

for any $(y_{t-1}, ..., y_{t-p_y}) \in Y^{(p)}$ and $\delta > 0$ when k is sufficiently large. Function $f_y(y_{t-1}, ..., y_{t-p_y})$ can thus be arbitrarily accurately approximated by the polynomial $T_y^k(0, y)$. A similar result holds for $f_x(y_{t-1}, ..., y_{t-p_x}, x_{t-1}, ..., x_{t-q_x})$ and $T_x^k(x, y)$. The null hypothesis H_{02} in (14) is a linear hypothesis in a linear system. From Assumptions A4, A5 and A6, and the fact that the approximation errors in (13) are negligible, it follows that the standard LM statistic for testing H_{02} is asymptotically χ^2-distributed with N_1 degrees of freedom when the null hypothesis holds, where N_1 is defined in (15).

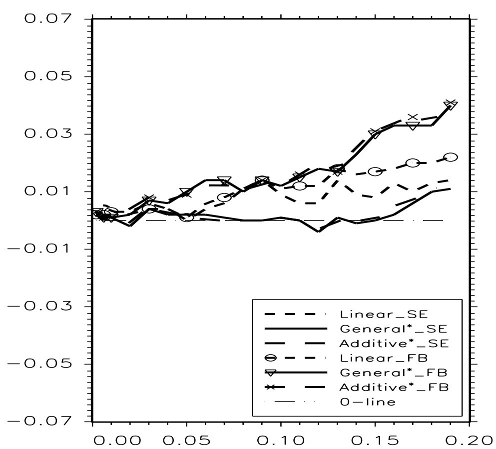


(a) x NGC y

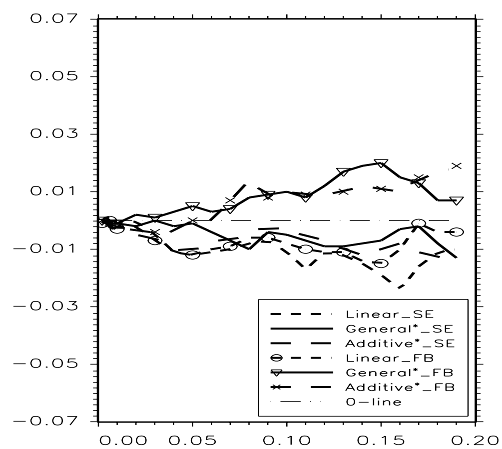


(b) y NGC x

Figure 1: Size discrepancy plots, data generated from system (21).

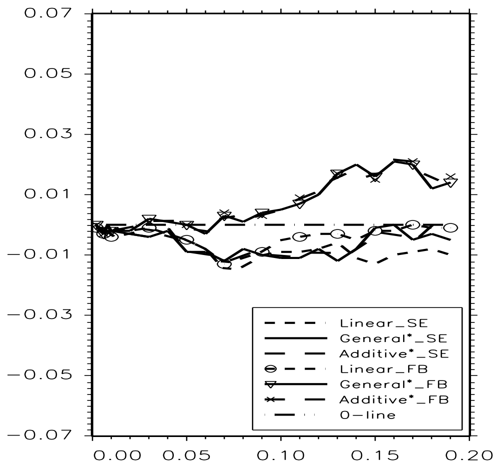


(a) x NGC y

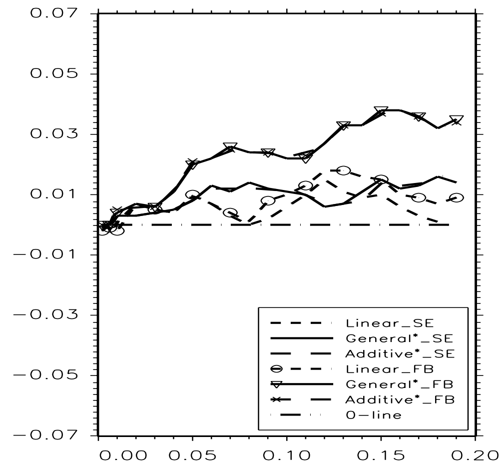


(b) y NGC x

Figure 2: Size discrepancy plots, data generated from system (22).

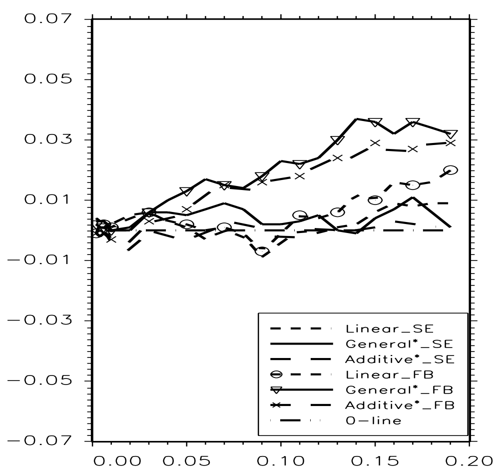


(a) x NGC y

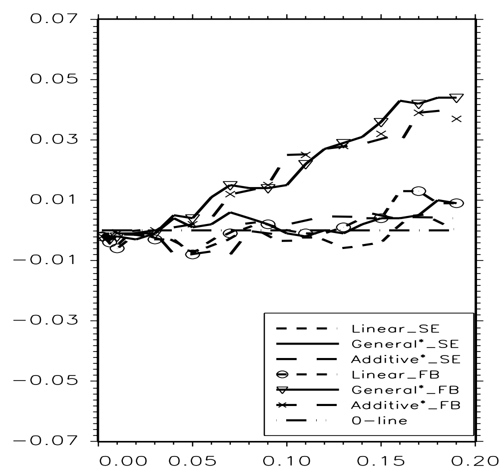


(b) y NGC x

Figure 3: Size discrepancy plots, data generated from system (23).

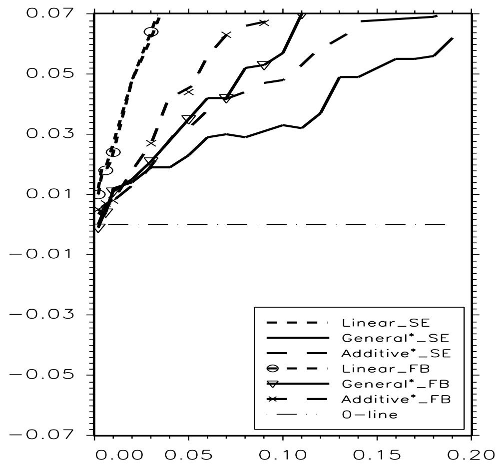


(a) x NGC y



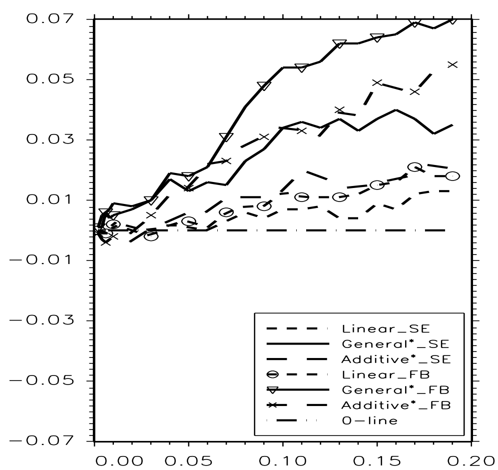
(b) y NGC x

Figure 4: Size discrepancy plots, data generated from system (24).

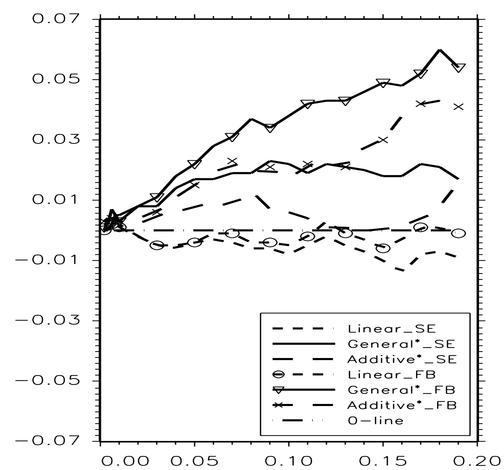


(a) x NGC y

Figure 5: Size discrepancy plots, data generated from system (25).

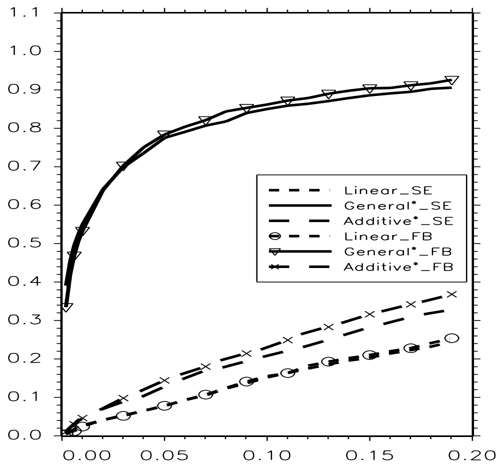


(a) x NGC y

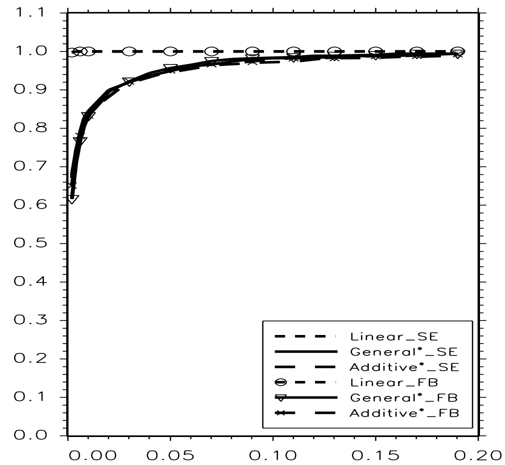


(b) y NGC x

Figure 6: Size discrepancy plots, data generated from system (26).

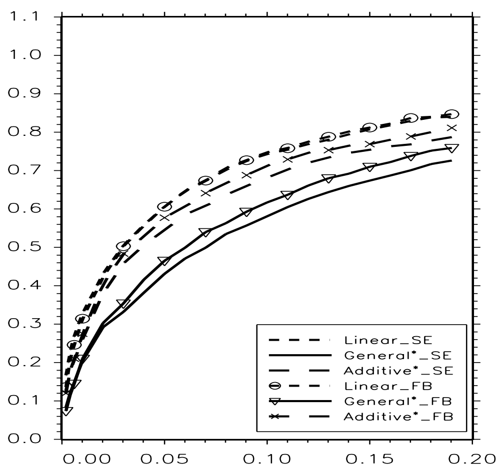


(a) x NGC y

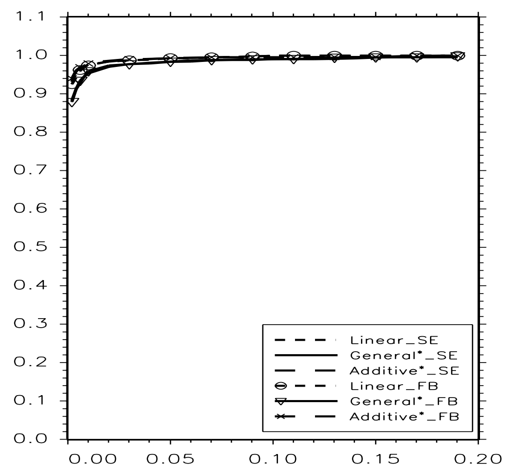


(b) y NGC x

Figure 7: Power-curves, data generated from system (27).

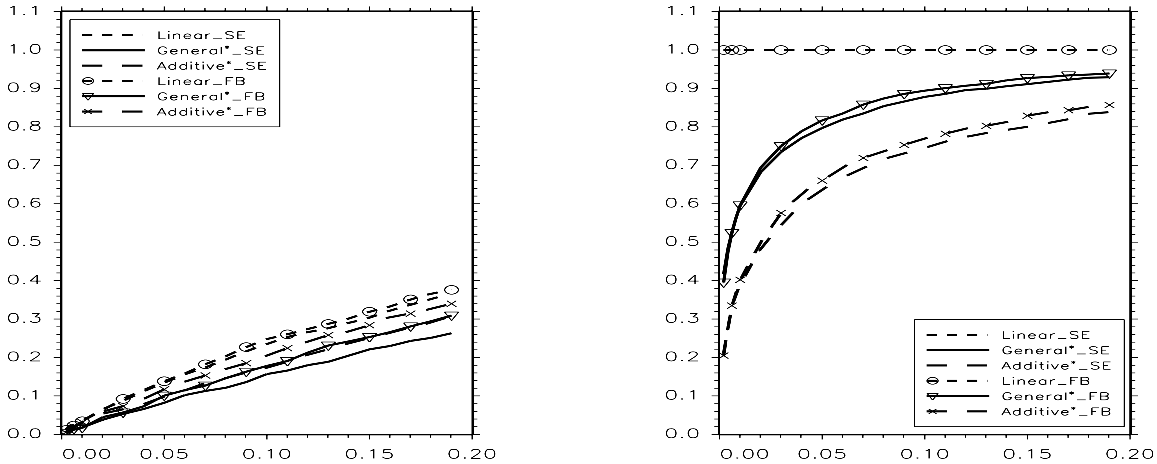


(a) x NGC y



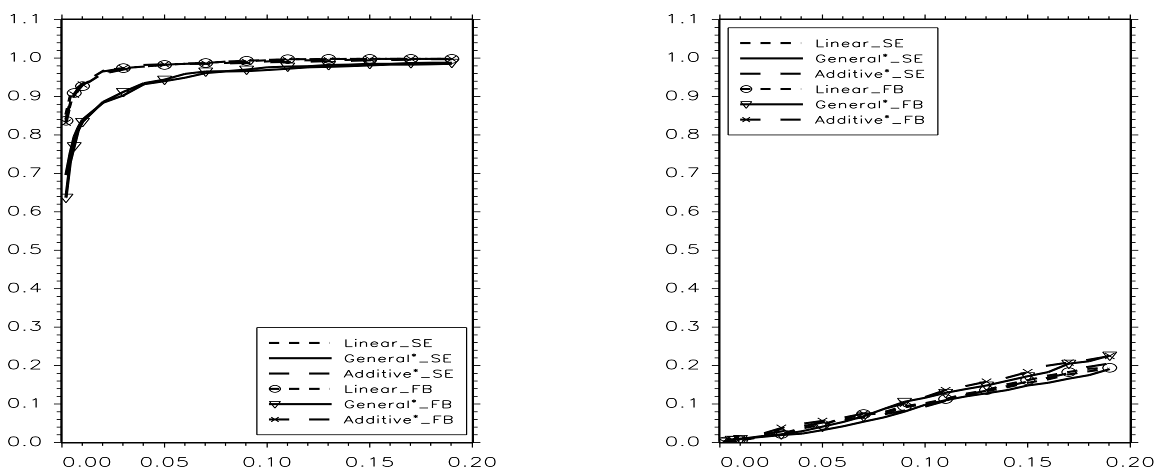
(b) y NGC x

Figure 8: Power-curves, data generated from system (28).



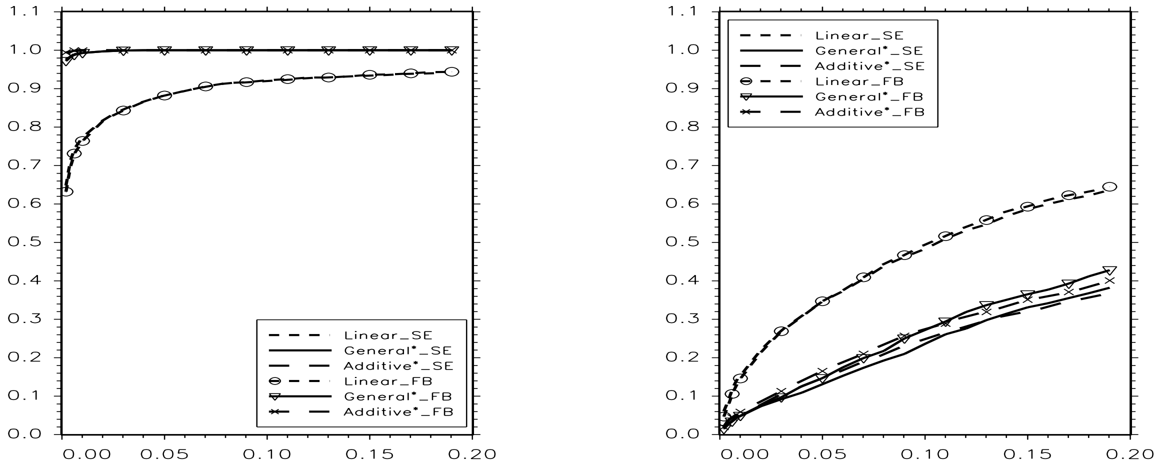
(a) x NGC y (b) y NGC x

Figure 9: Power-curves, data generated from system (29).



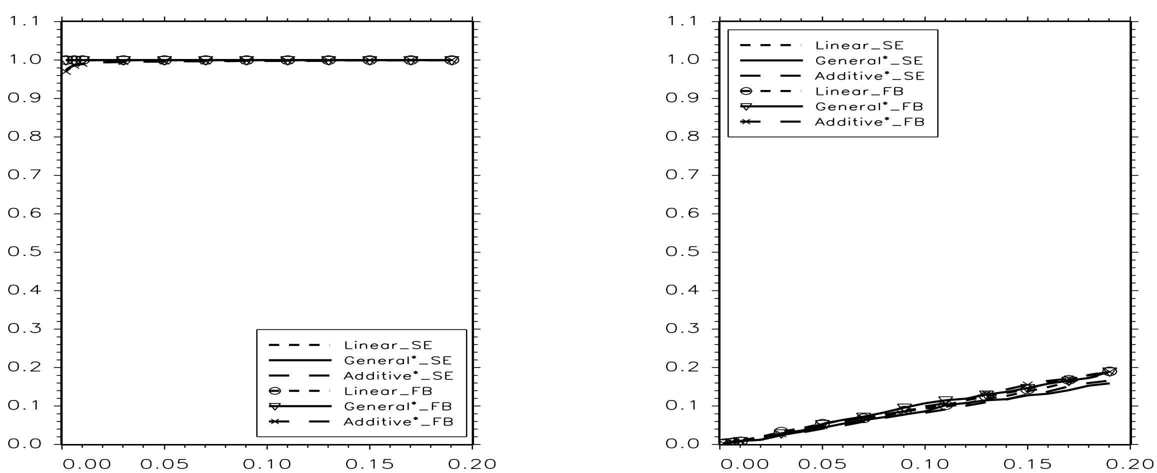
(a) x NGC y (b) y NGC x

Figure 10: Power-curves, data generated from system (30).



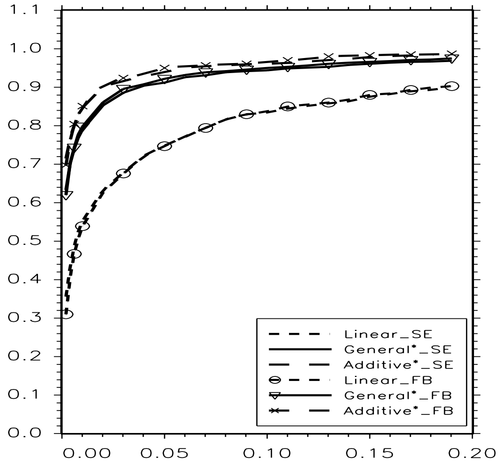
(a) x NGC y (b) y NGC x

Figure 11: Power-curves, data generated from system (31).

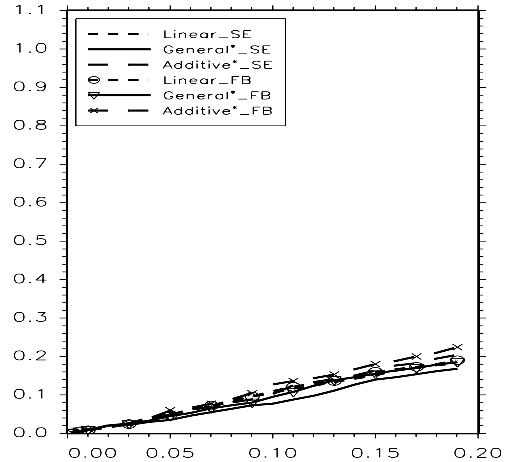


(a) x NGC y (b) y NGC x

Figure 12: Power-curves, data generated from system (32).

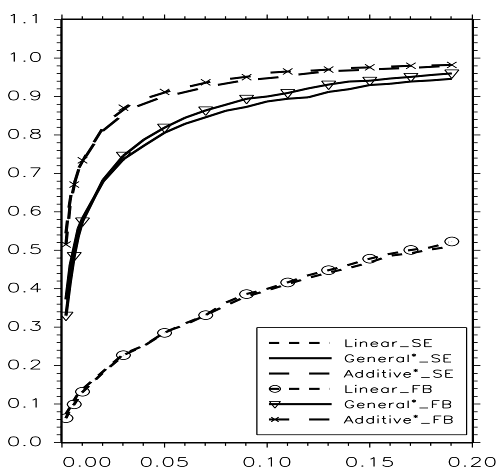


(a) x NGC y

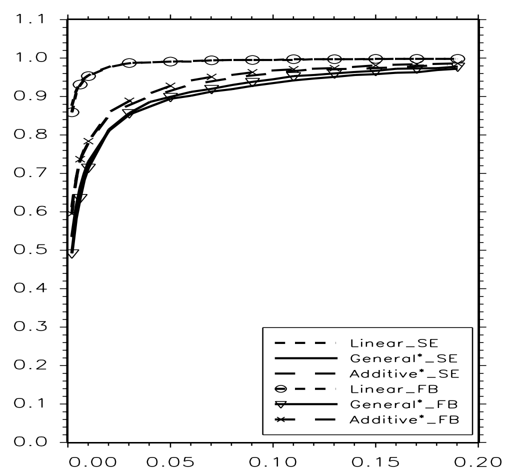


(b) y NGC x

Figure 13: Power-curves, data generated from system (33).

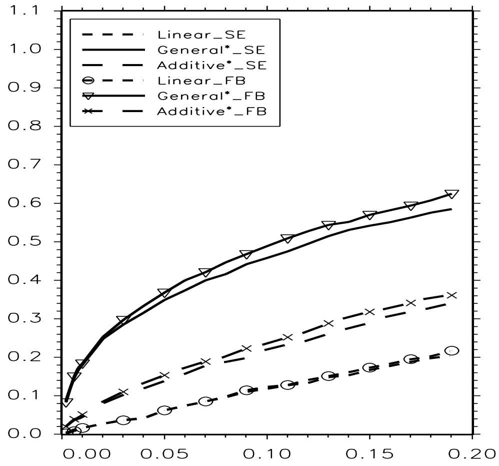


(a) x NGC y



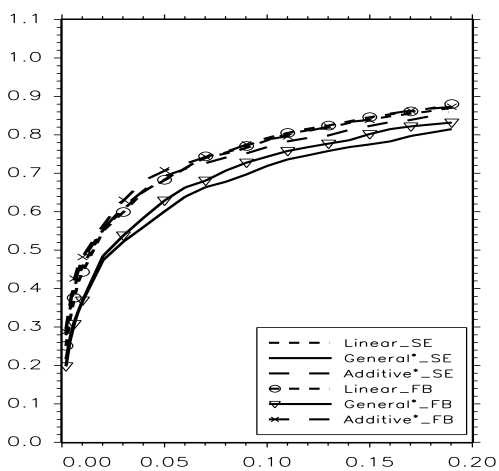
(b) y NGC x

Figure 14: Power-curves, data generated from system (34).



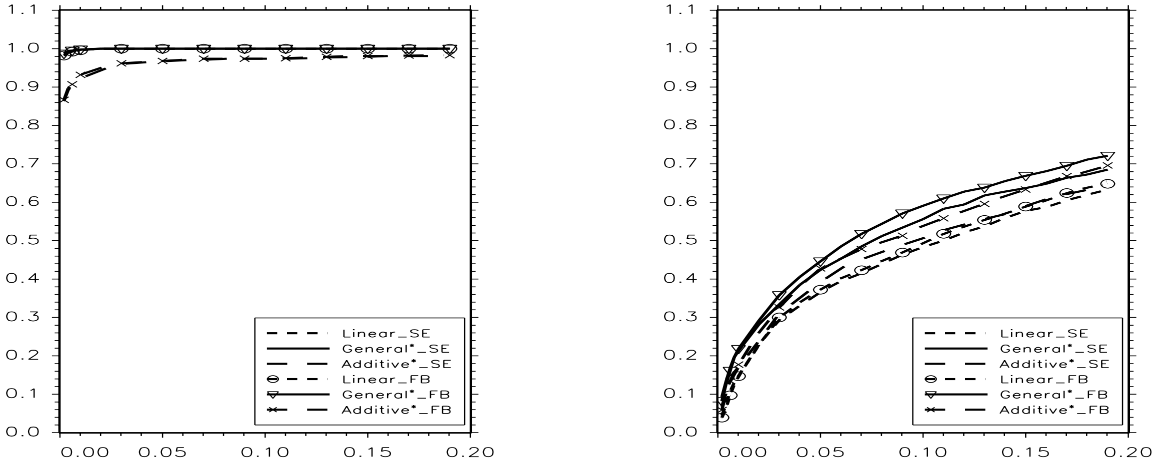
(a) x NGC y (b) y NGC x

Figure 15: Power-curves, data generated from system (35).



(a) x NGC y (b) y NGC x

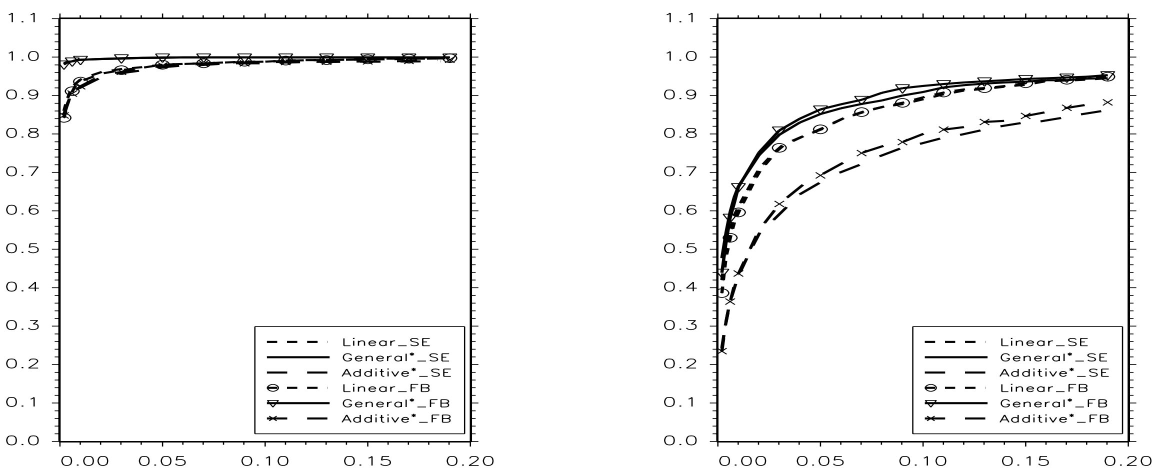
Figure 16: Power-curves, data generated from system eqrefeq:pow9.



(a) x NGC y

(b) y NGC x

Figure 17: Power-curves, data generated from system (37).



(a) x NGC y

(b) y NGC x

Figure 18: Power-curves, data generated from system (38).

Caused variable	q	IP	GDP	IMP	EXP	PROD	RW	INV	CONS	EMPL
	5									
	6									
IP	7									
$\hat{p} = 7$	8									
	9									
	0						*			
	5				*					
	6				*					
GDP	7				*					
$\hat{p} = 4$	8				*					
	9				*					
	0									
	5				***		*		*	
	6				***					
IMP	7				**					
$\hat{p} = 6$	8				**					
	9				***					
	0				***					
	5					*				
	6					*				*
EXP	7					*				*
$\hat{p} = 9$	8					*				*
	9					*			*	*
	0					*				*
	5	*					*		*	*
	6	*					*		*	*
PROD	7	**					*		*	*
$\hat{p} = 4$	8	**							*	*
	9	*					**			*
	0	*					**			**
	5			****	**	*		*	**	**
	6		*	****	***			*	**	*
RW	7		*	****	**			*	**	*
$\hat{p} = 4$	8		*	****	**				**	*
	9			****	**				*	*
	0		*	****	**				*	*
	5	*		*	*	*				
	6	***			*	**				
INV	7	***			*	*				
$\hat{p} = 4$	8	**			*	*				
	9	**				*				
	0	**				*				
	5	*		*	**					
	6				***					
CONS	7				***					
$\hat{p} = 9$	8				***					*
	9				**					*
	0				**					*
	5		**			*	****		**	
	6		**		*		****		**	
EMPL	7		**				****		**	
$\hat{p} = 4$	8		**		*		****		**	
	9		**		*		****		**	
	0		*		*		****		*	

Table 1: Results of linear GNC tests.

Caused variable	Causing variable									
	q	IP	GDP	IMP	EXP	PROD	RW	INV	CONS	EMPL
	5					***				***
	6					**		*		*
IP	7					**	*	*		*
$\hat{p} = 7$	8					**		*		*
	9					**	*	*		*
	0					**		*		**
	5				*					
	6									
GDP	7									
$\hat{p} = 4$	8									
	9									
	0									
	5				***		*			*
	6				****		*			**
IMP	7				****		*			**
$\hat{p} = 6$	8				****		*			**
	9	*	**		****		**			**
	0				****		**			**
	5					*			*	
	6					*				
EXP	7		*			**			*	
$\hat{p} = 9$	8					**			*	
	9			*		*			*	
	0			*		**				
	5	*					***		*	*
	6	*		*	*		***		*	*
PROD	7	**			*		**		*	*
$\hat{p} = 4$	8	*		*	*		**		*	
	9			*			***		*	
	0	*		**	*		**			**
	5			***	**					
	6			***	***					
RW	7			****	**					
$\hat{p} = 4$	8			***	**					
	9			***						
	0			***	**					
	5	*			*	*				
	6	***			*	**				
INV	7	**				**				
$\hat{p} = 4$	8	**				**				
	9	**				*				
	0	**				**				
	5				*					
	6				*					
CONS	7				*					
$\hat{p} = 9$	8			*	*					
	9									
	0			*	*					
	5	*	*				***			
	6						***			
EMPL	7						**			
$\hat{p} = 4$	8						***			
	9						**			
	0						**			

Table 2: Results of *Additive* GNC tests, Taylor expansion order $k = 2$.

Caused variable		Causing variable								
	q	IP	GDP	IMP	EXP	PROD	RW	INV	CONS	EMPL
	5					***				***
	6					**		*		**
IP	7					**	*	*		**
$\hat{p} = 7$	8					**				**
	9					**	*	*		***
	0					**	*	*		**
	5				*					
	6									
GDP	7									
$\hat{p} = 4$	8									
	9									
	0									
	5		****		****	*	*			**
	6		**		****	*			*	***
IMP	7	**	***		****	*				***
$\hat{p} = 6$	8	*	*		****	*	*			**
	9	**			****	*	*			**
	0	**			****	*	*			**
	5		*			**				
	6					**				*
EXP	7					**				*
$\hat{p} = 9$	8					***				
	9					**			*	*
	0	*				**			*	
	5	*		*	**		***		*	*
	6	*		*	**		**		*	*
PROD	7	**		*	**		**		*	*
$\hat{p} = 4$	8	*		**	*		**		*	*
	9	*		*			***			
	0	*		*	*		***			**
	5			***	****					
	6			**	****					
RW	7			***	*					
$\hat{p} = 4$	8			***	*					
	9			***	*					
	0			***	****					
	5	*				*				
	6	***				**				
INV	7	**				**				
$\hat{p} = 4$	8	**				**				
	9	**				*			*	
	0	**				**				
	5				*					
	6				*					
CONS	7				*					
$\hat{p} = 9$	8			*	*					
	9									
	0									
	5	*	*				***			
	6						***			
EMPL	7						**			
$\hat{p} = 4$	8						***			
	9						**			
	0						**			

Table 3: Results of *General* GNC tests, Taylor expansion order $k = 2$.

Research Papers 2008

CREATES
Center for Research in Econometric Analysis of Time Series

2008-06 Annastiina Silvennoinen and Timo Teräsvirta: Multivariate GARCH models. To appear in T. G. Andersen, R. A. Davis, J.-P. Kreiss and T. Mikosch, eds. Handbook of Financial Time Series. New York: Springer.

2008-07 Changli He, Annastiina Silvennoinen and Timo Teräsvirta: Parameterizing unconditional skewness in models for financial time series

2008-08 Cristina Amado and Timo Teräsvirta: Modelling Conditional and Unconditional Heteroskedasticity with Smoothly Time-Varying Structure

2008-09 Søren Johansen and Bent Nielsen: An analysis of the indicator saturation estimator as a robust regression estimator

2008-10 Peter Christoffersen, Kris Jacobs, Christian Dorion and Yintian Wang: Volatility Components, Affine Restrictions and Non-Normal Innovations

2008-11 Peter Christoffersen, Kris Jacobs, Chayawat Ornthanalai and Yintian Wang: Option Valuation with Long-run and Short-run Volatility Components

2008-12 Tom Engsted and Stig V. Møller: An iterated GMM procedure for estimating the Campbell-Cochrane habit formation model, with an application to Danish stock and bond returns

2008-13 Lars Stentoft: Option Pricing using Realized Volatility

2008-14 Jie Zhu: Pricing Volatility of Stock Returns with Volatile and Persistent Components

2008-15 Jie Zhu: Testing for Expected Return and Market Price of Risk in Chinese A-B Share Market: A Geometric Brownian Motion and Multivariate GARCH Model Approach

2008-16: Jie Zhu: FIEGARCH-M and and International Crises: A Cross-Country Analysis

2008-17: Almut E. D. Veraart: Inference for the jump part of quadratic variation of Itô semimartingales

2008-18: Michael Sørensen: Parametric inference for discretely sampled stochastic differential equations

2008-19: Anne Péguin-Feissolle, Birgit Strikholm and Timo Teräsvirta: Testing the Granger noncausality hypothesis in stationary nonlinear models of unknown functional form