

A note on the Hiemstra-Jones test for Granger non-causality

Cees Diks & Valentyn Panchenko

Center for Nonlinear Dynamics in Economics and Finance
Department of Economics, University of Amsterdam, Roetersstraat 11
1018 WB Amsterdam, The Netherlands

November 17, 2004

Abstract

We address a consistency problem in the commonly used nonparametric test for Granger causality developed by Hiemstra and Jones (1994). We show that the relationship tested is not implied by the null hypothesis of Granger non-causality. Monte Carlo simulations using processes satisfying the null hypothesis show that, for a given nominal size, the actual rejection rate may tend to one as the sample size increases. Our results imply that evidence for nonlinear Granger causality reported in the applied empirical literature should be re-interpreted.

Keywords: Granger causality; Hypothesis tests; Nonparametric tests

JEL classification: C12, C22, C52

1 Introduction

Consider a strictly stationary and weakly dependent bivariate time series process $\{(X_t, Y_t)\}$, $t \in \mathbf{Z}$. By definition, Y is strictly Granger (1969) causing X if the conditional distribution of X_t , given the past observations X_{t-1}, X_{t-2}, \dots and Y_{t-1}, Y_{t-2}, \dots , differs from the conditional distribution of X_t , given the past observations X_{t-1}, X_{t-2}, \dots only. Intuitively, Y is a Granger cause of X if adding past observations of Y to the information set increases the knowledge on the distribution of current values of X . Note that although the definition concerns conditional distributions given an infinite number of past observations, in practice tests are usually confined to finite orders in X and Y .

Several recent empirical studies report results for the Hiemstra and Jones (1994) test; a non-parametric test for Granger non-causality against general (linear and non-linear) alternatives. Evidence for causality is reported by e.g. Abhyankar (1998), Silvapulla and Moosa (1999), and Asimakopoulous *et al.* (2000). Okunev *et al.* (2002) report inefficiencies in Australian real estate and stock market prices. In this note we argue that the Hiemstra-Jones test does not provide a solid basis for conclusions of this type.

Initially, our interest was raised by the fact that, for some data sets, counter-intuitive results are obtained from the Hiemstra-Jones test and conventional tests of the same null hypothesis against linear Granger causality. Even if there is strong evidence for linear Granger causality, the Hiemstra-Jones test can fail to detect causality, or suggest that there is less causality than under the null hypothesis of no Granger causality (c.f. large negative values of test statistics reported in Brooks and Henry, 2000).

As an illustration we consider the test results presented in Table 1. We simulated 10000 observations from a bivariate data generating process (to be described in more detail later) with strong linear Granger causality from Y to X and applied linear and nonlinear Granger causality tests in both

Null Hypothesis:	Linear test		H-J test	
	F-stat	P-value	Z-stat	P-value
Y does not Granger Cause X	176.34	0.0000	-0.05	0.5204
X does not Granger Cause Y	2.64	0.1043	-6.79	1.0000

Table 1: Example test results for a simulated bivariate time series of length 10000.

directions for one lag. The linear test for Granger causality indicates strong evidence in support of the hypothesis that $\{Y_t\}$ Granger causes $\{X_t\}$. However, the Hiemstra-Jones test fails to detect the evident linear causal relation. Although this result may be counter-intuitive, it is not alarming in the sense that it merely indicates a lack of power of the Hiemstra-Jones test. However, the analysis presented in the following sections shows that this lack of power is in fact caused by an inconsistency which can lead to the much more severe problem of over-rejection.

2 The Hiemstra-Jones test

Following Hiemstra and Jones (1994), we let $F(X_t|\mathbf{I}_{t-1})$ denote the conditional probability distribution of X_t given the information set \mathbf{I}_{t-1} , which consists of an L_x -length lagged vector of X_t , say $\mathbf{X}_{t-L_x}^{L_x} \equiv (X_{t-L_x}, X_{t-L_x+1}, \dots, X_{t-1})$, and an L_y -length lagged vector of Y_t , say $\mathbf{Y}_{t-L_y}^{L_y} \equiv (Y_{t-L_y}, Y_{t-L_y+1}, \dots, Y_{t-1})$. Hiemstra and Jones consider testing, for a given pair of lags L_x and L_y , the following relationship:

$$H_0 : \quad F(X_t|\mathbf{I}_{t-1}) = F(X_t|\mathbf{I}_{t-1} - \mathbf{Y}_{t-L_y}^{L_y}). \quad (1)$$

That is, the null hypothesis of interest states that taking the vector of past Y -values out of the information set does not affect the distribution of current X -values.

Adopting the notation used by Hiemstra and Jones, we denote the m -length lead vector of X_t by \mathbf{X}_t^m , so that we can summarize the vectors defined so far, for $t \in \mathbf{Z}$, as:

$$\begin{aligned} \mathbf{X}_t^m &= (X_t, X_{t+1}, \dots, X_{t+m-1}), & m &= 1, 2, \dots \\ \mathbf{X}_{t-L_x}^{L_x} &= (X_{t-L_x}, X_{t-L_x+1}, \dots, X_{t-1}), & L_x &= 1, 2, \dots \\ \mathbf{Y}_{t-L_y}^{L_y} &= (Y_{t-L_y}, Y_{t-L_y+1}, \dots, Y_{t-1}), & L_y &= 1, 2, \dots \end{aligned} \quad (2)$$

A crucial claim made by Hiemstra and Jones (1994) without proof, states that the null hypothesis given in Eq. (1) implies, for all $\epsilon > 0$:

$$\begin{aligned} P \left(\|\mathbf{X}_t^m - \mathbf{X}_s^m\| < \epsilon \mid \|\mathbf{X}_{t-L_x}^{L_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon, \|\mathbf{Y}_{t-L_y}^{L_y} - \mathbf{Y}_{s-L_y}^{L_y}\| < \epsilon \right) \\ = P \left(\|\mathbf{X}_t^m - \mathbf{X}_s^m\| < \epsilon \mid \|\mathbf{X}_{t-L_x}^{L_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon \right), \end{aligned} \quad (3)$$

where $P(A|B)$ denotes the conditional probability of A given B , and $\|\cdot\|$ the supremum norm, which for a d -dimensional vector $x = (x_1, \dots, x_d)^T$ is given by $\|x\| = \sup_{i=1}^d |x_i|$. Equation (3) states that the conditional probability that two arbitrary m -length lead vectors of $\{X_t\}$ are within distance ϵ , given that the corresponding lagged L_x -length lag vectors of $\{X_t\}$ are ϵ -close, is the same as when in addition one also conditions on the L_y -length lag vectors of $\{Y_t\}$ being ϵ -close.

For an observed bivariate time series $\{(X_t, Y_t)\}$, $t = 1, \dots, T$, the Hiemstra-Jones test consists of choosing a value for ϵ (typical values are between 0.5 and 1.5 after normalizing the time series to

unit variance), and testing (3) by estimating the conditional probabilities as ratios of unconditional probabilities. Using the notation

$$\begin{aligned}
C1(m + L_x, L_y, \epsilon) &= P(\|\mathbf{X}_{t-L_x}^{m+L_x} - \mathbf{X}_{s-L_x}^{m+L_x}\| < \epsilon, \|\mathbf{Y}_{t-L_y}^{L_y} - \mathbf{Y}_{s-L_y}^{L_y}\| < \epsilon), \\
C2(L_x, L_y, \epsilon) &= P(\|\mathbf{X}_{t-L_x}^{L_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon, \|\mathbf{Y}_{t-L_y}^{L_y} - \mathbf{Y}_{s-L_y}^{L_y}\| < \epsilon), \\
C3(m + L_x, \epsilon) &= P(\|\mathbf{X}_{t-L_x}^{m+L_x} - \mathbf{X}_{s-L_x}^{m+L_x}\| < \epsilon), \\
C4(L_x, \epsilon) &= P(\|\mathbf{X}_{t-L_x}^{L_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon),
\end{aligned} \tag{4}$$

equation (3) can be expressed as

$$\frac{C1(m + L_x, L_y, \epsilon)}{C2(L_x, L_y, \epsilon)} = \frac{C3(m + L_x, \epsilon)}{C4(L_x, \epsilon)}. \tag{5}$$

Upon denoting the U-statistics estimators of the probabilities in (4) with an additional index n , where $n = T + 1 - m - \max(L_x, L_y)$, Hiemstra and Jones show that, under (3):

$$\sqrt{n} \left(\frac{C1(m + L_x, L_y, \epsilon, n)}{C2(L_x, L_y, \epsilon, n)} - \frac{C3(m + L_x, \epsilon, n)}{C4(L_x, \epsilon, n)} \right) \underset{a}{\sim} N(0, \sigma^2(m, L_x, L_y, \epsilon)), \tag{6}$$

with $\sigma^2(m, L_x, L_y, \epsilon)$ as given in their appendix. One-sided critical values are used, based on this asymptotic result, rejecting when the observed value of the test statistic in (6) is too large.

3 A counter-example

In this section we show that equation (3) is not, in general, implied by the null hypothesis given in (1). The left hand side of equation (3) under the null can be either smaller or larger than the right hand side. To emphasize the relevance to econometrics and financial time series analysis, we focus on a class of bivariate time series processes with conditional heteroskedasticity. The starting point is a first order process $\{X_t\}$ with conditional heteroskedasticity

$$X_t \sim N(0, g(X_{t-1})), \tag{7}$$

where $g(\cdot)$ is some positive function, such that $\{X_t\}$ is stationary and ergodic (e.g. if $g(X_{t-1}) = a + bX_{t-1}^2$ with $a > 0$ and $0 < b < 1$, X_t is a stable ARCH(1) process). Next consider a process $\{Y_t\}$, which is instantaneously driven by $\{X_t\}$ via:

$$Y_{t-1} \sim N(0, h(X_{t-1})), \tag{8}$$

where $h(\cdot)$ is again a non-negative function, and where Y_{t-1} is taken to be conditionally independent of X_t given X_{t-1} . In the particular case where $h(s) = g(s)$ for all s , X_t and Y_{t-1} are identically distributed, conditionally on X_{t-1} . Clearly, Y is not Granger-causing X . A single lagged value X_{t-1} fully determines the distribution of X_t , so that equation (1) holds for all $L_x, L_y = 1, 2, \dots$

Next we show that $g(\cdot)$ and $h(\cdot)$ can be chosen in such a way that equation (3) does not hold for any positive value of ϵ . A convenient first step is to rephrase equation (3) as

$$\begin{aligned}
&P \left(\|\mathbf{X}_t^m - \mathbf{X}_s^m\| < \epsilon, \|\mathbf{Y}_{t-L_y}^{L_y} - \mathbf{Y}_{s-L_y}^{L_y}\| < \epsilon \mid \|\mathbf{X}_{t-L_x}^{L_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon \right) \\
&= P \left(\|\mathbf{X}_t^m - \mathbf{X}_s^m\| < \epsilon \mid \|\mathbf{X}_{t-L_x}^{L_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon \right) \\
&\quad \times P \left(\|\mathbf{Y}_{t-L_y}^{L_y} - \mathbf{Y}_{s-L_y}^{L_y}\| < \epsilon \mid \|\mathbf{X}_{t-L_x}^{L_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon \right),
\end{aligned} \tag{9}$$

obtained by rewriting equation (3), which is of the form $P(A|B, C) = P(A|B)$, in the form $P(A, C|B) = P(A|B)P(C|B)$. As Hiemstra and Jones we focus on the case $m = 1$ (one-step lead vectors of $\{X_t\}$). In that case we find for $L_y = 1$:

$$\begin{aligned} & P\left(|X_t - X_s| < \epsilon, |Y_{t-1} - Y_{s-1}| < \epsilon \mid \|\mathbf{X}_{t-l_x}^{l_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon\right) \\ &= P\left(|X_t - X_s| < \epsilon \mid \|\mathbf{X}_{t-l_x}^{l_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon\right) \\ &\quad \times P\left(|Y_{t-1} - Y_{s-1}| < \epsilon \mid \|\mathbf{X}_{t-l_x}^{l_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon\right). \end{aligned} \quad (10)$$

The difference between the left and right hand sides can be expressed as a conditional covariance

$$D = \text{Cov}\left(I(|X_t - X_s| < \epsilon), I(|Y_{t-1} - Y_{s-1}| < \epsilon) \mid \|\mathbf{X}_{t-l_x}^{l_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon\right), \quad (11)$$

where $I(\cdot)$ denotes the indicator function which is one if the inequality in its argument holds and zero otherwise.

Although for given values of X_{t-1} and X_{s-1} , $X_t - X_s$ and $Y_{t-1} - Y_{s-1}$ are independent normally distributed random variables (with variances $V = g(X_{t-1}) + g(X_{s-1})$ and $W = h(X_{t-1}) + h(X_{s-1})$, respectively) the condition in equation (11) does not fix X_{t-1} and X_{s-1} , so that the covariance need not be zero. We obtain

$$D = \text{Cov}\left(I(|Z_1|\sqrt{V} < \epsilon), I(|Z_2|\sqrt{W} < \epsilon)\right), \quad (12)$$

where Z_1 and Z_2 are standard normal random variables, independent of each other and (V, W) , and V and W are the conditional variances of $X_t - X_s$ and $Y_{t-1} - Y_{s-1}$, given $\|\mathbf{X}_{t-l_x}^{l_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon$, which are jointly distributed according to

$$(V, W) \sim (g(X_{t-1}) + g(X_{s-1}), h(X_{t-1}) + h(X_{s-1})) \mid \|\mathbf{X}_{t-l_x}^{l_x} - \mathbf{X}_{s-L_x}^{L_x}\| < \epsilon. \quad (13)$$

Since Z_1 and Z_2 are independent, and independent of (V, W) , the expectations with respect to those variables can be taken. If we define

$$r(s) = E(I(|Z| < s)) = P(|Z| < s), \quad \text{for } Z \sim N(0, 1), \quad (14)$$

we obtain

$$D = \text{Cov}\left(r(\epsilon/\sqrt{V}), r(\epsilon/\sqrt{W})\right). \quad (15)$$

Depending on the joint distribution of V and W , D can be either negative, zero, or positive. The most problematic case is $D > 0$, since the one-sided Hiemstra-Jones test will then tend to over-reject. Clearly, if either V or W is degenerate (i.e. with probability one takes only one specific value), the covariance is zero and $D = 0$. The case $D > 0$ thus requires V and W to be non-degenerate random variables. Let us focus on V first. The fact that V can have a non-degenerate distribution follows from the existence of stationary ARCH(1) processes with time varying conditional variance. If for such non-degenerate V we define W in such a way that it is positively correlated with V , then $D > 0$. An obvious example would be to take $g(s) = h(s)$ for all s , which implies $W = V$. In that case one finds $D = \text{Var}(r(\epsilon/\sqrt{V})) > 0$.

Further analytic results, which are beyond the scope of this paper, indicate that also for processes of a different form than that in equations (7) and (8) but which also satisfy the null hypothesis, D typically is nonzero. These results also suggest some ways of reducing the bias, which we hope to prove useful in future work on alternative tests. The fact that the sizes reported in the bootstrap study

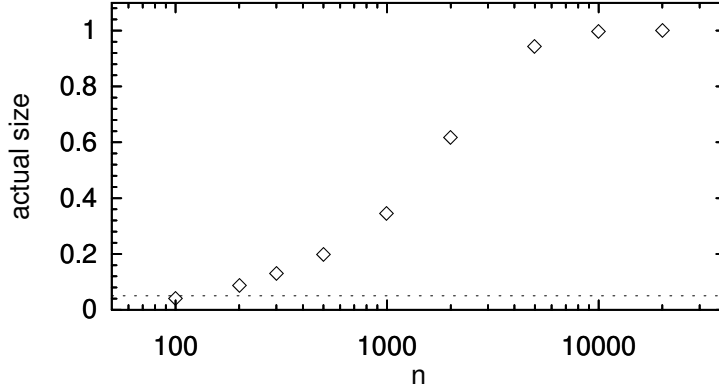


Figure 1: Simulated size of the Hiemstra-Jones test ($\epsilon = 1$, $L_x = L_y = 1$) for the bivariate ARCH process given in equation (17) as a function of the time series length (nominal size 0.05). Number of realizations: 1000 for $n < 10000$, and 500, 100 for $n = 10000$ and 20000 respectively.

by Diks and DeGoede (2001) were close to nominal is related to the relatively small sample sizes used there.

The data generating process for the time series used to generate Table 1 was:

$$\begin{aligned} X_t &\sim N(0.2Y_{t-1}, 1 + 0.4X_{t-1}^2) \\ Y_{t-1} &\sim N(0, (1 + 0.4X_{t-1}^2)^{-1}). \end{aligned} \quad (16)$$

The resulting time series processes are stationary and exhibit conditional heteroskedasticity. Clearly, $\{Y_t\}$ linearly Granger causes $\{X_t\}$. At the same time there is contemporaneous nonlinear dependence between $\{Y_t\}$ and $\{X_t\}$. There is a negative bias in the Hiemstra-Jones test statistic introduced by the negative dependence between the conditional variance of Y_{t-1} and X_t given X_{t-1} . This destroys power in the Granger non-causality test from Y to X and leads to smaller test statistics than expected under the null in the test of Granger non-causality from X to Y .

Simulations

To illustrate the effect of a positive covariance in equation (15) we imposed V and W to be equal (see eq. 13) by taking $g(X_{t-1}) = h(X_{t-1}) = 1 + 0.4X_{t-1}^2$. We simulated the resulting bivariate ARCH process,

$$\begin{aligned} X_t &\sim N(0, 1 + 0.4X_{t-1}^2) \\ Y_{t-1} &\sim N(0, 1 + 0.4X_{t-1}^2), \end{aligned} \quad (17)$$

and calculated the rejection rates (at nominal size 0.05) of the Hiemstra-Jones test for the null hypothesis that Y does not Granger cause X .

Figure 1 shows the rejection rates found as a function of the time series length n . The actual size of the test is close to the nominal size of 0.05 only for the short time series of 100. The size increases with the length n of the time series, and is close to one already for time series of length 5000.

4 Case study

In practice the test is usually applied after filtering out seasonalities, linear structure, and (G)ARCH structure. Although this may lead to smaller rejection rates due to whitening of the data, it does not

Procedure	$n = 1000$		$n = 10000$	
	Z-stat	P-value	Z-stat	P-value
Raw data	3.183796	0.0007	7.893391	0.0000
GARCH(1,1)	1.689628	0.0455	5.823490	2.9E-9
GARCH(1,1) & AR(1)	2.362784	0.0091	4.932293	4.1E-7
Correct model	1.323609	0.0928	0.082326	0.4672

Table 2: Results of the control for the data structure in case study

affect our conclusion in typical cases where the model specification is not known to be correct. To illustrate this point we mimic a typical empirical study relying on the Hiemstra-Jones test by investigating an artificial bivariate process of the form

$$\begin{aligned} X_t &\sim N(-2X_{t-1}e^{-X_{t-1}^2}, 1 + 0.4X_{t-1}^2) \\ Y_{t-1} &\sim N(-2X_{t-1}e^{-X_{t-1}^2}, 1 + 0.4X_{t-1}^2). \end{aligned} \quad (18)$$

The process (18) satisfies the null hypothesis that Y does not Granger cause X . The X -series exhibits nonlinear AR(1) dependence in the mean and ARCH(1) structure, while Y is instantaneously driven by X through the mean and variance. Time series of 1000 and 10000 observations were considered for the study.

As mentioned above, a researcher usually controls the data for some structure. The univariate GARCH(1,1) model is a popular choice in financial time series. For both series we apply the GARCH(1,1) filter with two different mean specifications. First, we consider a simple model with constant mean. Thereafter, also AR(1) structure typical in financial studies is included in the mean equation. Table 4 summarizes the results of the Hiemstra-Jones tests after the above procedures. The results for the raw data and the residuals of the correctly specified model are included for reference. The test on the raw data strongly rejects the null hypothesis while it holds, as a result of the bias. Controlling for GARCH(1,1) with constant mean reduces the bias, but because of misspecification does not remove it completely. Adding AR(1) to the mean equation worsens the bias compared to the former procedure when the series length is 1000. This suggests that removing (G)ARCH and AR structure without knowledge of the correct model class may not correct the test and consequently, produce unreliable results. The test on the residuals of the correctly specified model leads to an anticipated result since the residual series by construction are practically independent, in which case equation (3) holds. In that sense the Hiemstra-Jones test performed on the residuals may be considered as a model specification test.

5 Conclusions

The analytic and numerical evidence presented in this note clearly show that equation (3), which is the relationship tested in the Hiemstra-Jones test, is not generally compatible with the null hypothesis stated in equation (1). This indicates that rejections of the null hypothesis reported in the empirical literature may be spurious. A simulated empirical study shows that our main conclusion remains valid also for studies which corrected for AR and (G)ARCH structure.

One might still argue, correctly, that the Hiemstra-Jones test is a valid test of the relationship given in equation (3). In fact one might even go one step further and take equation (3) as a definition of Granger causality, which is exactly the approach taken in the original test by Baek and Brock (1992). Although one can indeed test (3) using the Hiemstra-Jones test, the interpretation involves

some subtleties. A problem with this approach is that it is hard to find out in detail exactly which subclasses of data generating processes satisfy the null hypothesis. Although it is easy to give some sufficient conditions for (3) to hold for all ϵ (for example, $\{X_t\}$ and $\{Y_t\}$ being independent) it is surprisingly difficult to formulate necessary conditions in terms of the data generating process.

Finally, our results suggest that, since for $\{X_t\}$ and $\{Y_t\}$ independent (3) holds for all ϵ , the Hiemstra-Jones test can still be used as a model specification test by applying it to the residuals of an estimated model for the data generating process. However, in that case the Hiemstra-Jones test is used as a test for independence rather than conditional independence.

References

- Abhyankar, A. (1998). Linear and non-linear granger causality: Evidence from the U.K. stock index futures market. *Journal of Futures Markets*, **18**, 519–540.
- Asimakopoulou, I., Ayling, D. and Mahmood, W. M. (2000). Non-linear granger causality in the currency futures returns. *Economics Letters*, **68**, 25–30.
- Baek, E. and Brock, W. (1992). A general test for granger causality: Bivariate model. Technical Report. Iowa State University and University of Wisconsin, Madison.
- Brooks, C. and Henry, O. T. (2000). Linear and non-linear transmission of equity return volatility: evidence from the US, Japan and Australia. *Economic Modelling*, **17**, 497–513.
- Diks, C. and DeGoede, J. (2001). A general nonparametric bootstrap test for granger causality. In *Global Analysis of Dynamical Systems* (eds H. W. Broer, B. Krauskopf and G. Vegter), pp. 391–403. IoP Publishing, Bristol.
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, **37**, 424–438.
- Hiemstra, C. and Jones, J. D. (1994). Testing for linear and nonlinear Granger causality in the stock price-volume relation. *Journal of Finance*, **49**, number 5, 1639–1664.
- Okunev, J., Wilson, P. and Zurbrugg, R. (2002). Relationships between Australian real estate and stock market prices - a case of market inefficiency. *Journal of Forecasting*, **21**, number 3, 181–192.
- Silvapulla, P. and Moosa, I. A. (1999). The relationship between spot and futures prices: Evidence from the crude oil market. *Journal of Futures Markets*, **19**, 157–193.