I N S T I T U T D E S T A T I S T I Q U E
B I O S T A T I S T I Q U E E T
S C I E N C E S A C T U A R I E L L E S
(I S B A)

UNIVERSITÉ CATHOLIQUE DE LOUVAIN

D I S C U S S I O N
P A P E R

2012/09

Nonparametric Estimation and Inference for Granger Causality Measures

TAAMOUTI, A., BOUEZMARNI, T. and A. EL GHOUCH

Nonparametric Estimation and Inference for Granger Causality Measures*

Abderrahim Taamouti† Taoufik Bouezmarni‡ Anouar El Ghouch §

March 29, 2012

ABSTRACT

We propose a nonparametric estimator and a nonparametric test for *Granger causality measures* that *quantify* linear and nonlinear Granger causality in distribution between random variables. We first show how to write the Granger causality measures in terms of *copula densities*. We suggest a consistent estimator for these causality measures based on nonparametric estimators of copula densities. Further, we prove that the nonparametric estimators are asymptotically normally distributed and we discuss the validity of a local smoothed bootstrap that we use in finite sample settings to compute a bootstrap bias-corrected estimator and test for our causality measures. A simulation study reveals that the bias-corrected bootstrap estimator of causality measures behaves well and the corresponding test has quite good finite sample size and power properties for a variety of typical data generating processes and different sample sizes. Finally, we illustrate the practical relevance of nonparametric causality measures by quantifying the Granger causality between S&P500 Index returns and many exchange rates (US/Canada, US/UK and US/Japen exchange rates).

JEL Classification: C12; C14; C15; C19; G1; G12; E3; E4.

Key words: Causality measures; Nonparametric estimation; time series; copulas; Bernstein copula density; local bootstrap; conditional distribution function; stock returns.

*We would like to thank the seminar participants in Humboldt University of Berlin and Institut de statistique, Biostatistique et science actuarielles (ISBA) at Université catholique de Louvain for excellent comments that improved this paper. Financial support from the Spanish Ministry of Education through grants SEJ 2007-63098 is also acknowledged.

A. El Ghouch acknowledges financial support from IAP research network P6/03 of the Belgian Government (Belgian Science Policy), and from the contract 'Projet d'Actions de Recherche Concertées' (ARC) 11/16-039 of the 'Communauté française de Belgique', granted by the 'Académie universitaire Louvain'.

†Economics Department, Universidad Carlos III de Madrid, España; E-mail: ataamout@eco.uc3m.es.

‡Département de mathématiques, Université de Sherbrooke, Québec; E-mail: Taoufik.Bouezmarni@usherbrooke.ca.

§ISBA, Université Catholique de Louvain, Louvain-la-Neuve, Belgium; E-mail: Anouar.Elghouch@uclouvain.be.

1 Introduction

As pointed out by Geweke (1982), much research has been devoted to building and applying tests of non-causality. However, once we have concluded that a "causal relation" (in the sense of Granger) is present, it is usually important to assess the strength of the relationship [see also Dufour and Taamouti (2010)]. Few papers have been proposed to measure the causality between random variables. Further, though the concept of causality is naturally defined in terms of conditional distributions, the estimation of the existing causality measures has been done using parametric mean regression models in which the causal link between the variables of interest is linear. Consequently, one simply cannot use the latter estimated measures to quantify the strength of high-order moment causalities and nonlinear causalities. The present paper aims to propose a nonparametric estimator and a nonparametric test for Granger causality measures. The proposed approach is model-free and allows to quantify linear and nonlinear and low and high-order moments causalities.

The concept of causality introduced by Wiener (1956) and Granger (1969) constitutes a basic notion for studying dynamic relationships between time series. This concept is defined in terms of predictability at horizon one of a (vector) variable Y from its own past, the past of another (vector) variable Z, and possibly a vector X of auxiliary variables. The theory of Wiener-Granger causality has generated a considerable literature; for reviews, see Pierce and Haugh (1977), Newbold (1982), Geweke (1984a), Lütkepohl (1991), Boudjellaba, Dufour, and Roy (1992), Boudjellaba, Dufour, and Roy (1994), Gouriéroux and Monfort (1997, Chapter 10), and Dufour and Taamouti (2010).

Wiener-Granger analysis distinguishes between three basic types of causality: from Y to Z, from Z to Y, and instantaneous causality. In practice, it is possible that all three causality relations coexist, hence the importance of finding means to quantify their degree. Unfortunately, causality tests fail to accomplish this task, because they only provide evidence on the presence of causality. A large effect may not be statistically significant (at a given level), and a statistically significant effect may not be "large" from an economic viewpoint (or more generally from the viewpoint of the subject at hand) or relevant for decision making. Thus, as emphasized by McCloskey and Ziliak (1996), it is crucial to distinguish between the numerical value of a parameter and its statistical significance.

In studying Wiener-Granger causality, predictability is the central issue. So, beyond accepting or rejecting non-causality hypotheses – which state that certain variables do not help forecasting other variables – we wish to assess the magnitude of the forecast improvement, where the latter is defined in terms of some loss function (Kullback distance). Even if the hypothesis of no improvement (non-causality) cannot be rejected from looking at the available data (for example, because the sample size or the structure of the process does allow for high test power), sizeable improvements may remain consistent with the same data. Or, by contrast, a statistically significant improvement – which may easily be produced by a large data set - may not be relevant from a practical viewpoint.

The topic of measuring the causality has attracted much less attention. Geweke (1982) and Geweke (1984b) introduced measures of causality based on mean-square forecast errors. Gouriéroux, Monfort, and Renault (1987) proposed causality measures based on the Kullback information criterion and provided a *parametric* estimation for their measures. Polasek (1994) showed how causality measures can be computed using the Akaike Information Criterion (AIC). Polasek (2002) also introduced new causality measures in the context of univariate and multivariate ARCH models and their extensions based on a Bayesian approach. Finally, Dufour and Taamouti (2010) proposed a short and long run causality measures. The estimation of the above causality measures has been done in the context of parametric mean regression models; or under normality assumption. However, it is well known that the misspecification of the parametric model may affect the structure of the causality between the random variables of interest. Further, the dependence in the mean-regression model is only due to the mean dependence, thus this ignores the dependence in high-order moments. Finally, as shown in many theoretical and empirical papers, several "causal relations" are nonlinear; see for example Gabaix, Gopikrishnan, V. Plerou, and Stanley (2003), Hiemstra and Jones (1994), Bouezmarni, Rombouts, and Taamouti (2011) and Bouezmarni, Roy, and Taamouti (2010) among others. Consequently, the above estimation methods of causality measures can not be applied to quantify, for example, the strength of nonlinear causalities.

Here we propose a nonparametric estimator for Granger causality measures. The latter are initially defined in terms of conditional densities, thus they can capture and quantify linear and nonlinear causalities and the causalities due to both low and high-order moments. The nonparametric estimation method is model-free, and therefore it does not require the specification of the model relating the variable of interest. We first rewrite the theoretical Granger causality measures in terms of copula densities, which allow us to disentangle the dependence structure from the marginal distributions. Thereafter, the causality measures are estimated by replacing the unknown copula densities by their nonparametric estimates, which give full weight to the data. We use the Bernstein copula density to estimate nonparametrically the copula densities. For i.i.d. data, Sancetta and Satchell (2004) show that under some regularity conditions, any copula can be represented by a Bernstein copula. Bouezmarni, Rombouts, and Taamouti (2010) provide asymptotic properties of the Bernstein copula density estimator using dependent data. The nonparametric Bernstein copula density estimates are guaranteed to be non-negative and therefore we avoid potential problems with the function of measurement, in this case the logarithmic function in the Kullback *distance*. Furthermore, there is no boundary bias problem when we use the copula density estimator, because, by smoothing with beta densities, the Bernstein copula density does not assign weight outside its support.

To construct tests for Granger causality measures, we show that the nonparametric estimator of the latter is asymptotically normally distributed. To achieve this result, we subtract some bias terms from the Kullback distance between the copula densities and then rescale by the proper variance. Furthermore, we discuss the validity of the local smoothed bootstrap that we use in finite sample settings to compute a bootstrap bias-corrected estimator and test for the Granger causality measures. A simulation study reveals that the bootstrap bias-corrected estimator of causality measures is quite good and the test has good finite sample size and power properties for a variety of typical data generating processes and different sample sizes. The empirical importance of measuring high-order and nonlinear causalities is illustrated. Thus, in an empirical application we quantify the causality between S&P500 Index returns and many exchange rates (US/Canada, US/UK and US/Japen exchange rates). We find that both exchange rates and stock prices could have a significant impact on each other. We also find that the impacts of stock returns on exchange rates are much stronger than the impacts of exchange rates on stock returns.

The plan of the paper is as follows. Section 2 provides the motivation for considering a nonparametric causality measures. Sections 3 and 4 present the general theoretical framework which underlie the definitions of causality measures using probability and copula density functions. In Section 5, we introduce a consistent nonparametric estimator of causality measures, in the context of univariate random variables, based on Bernstein polynomial, and we establish the asymptotic distribution of the nonparametric estimator and discuss the asymptotic validity of a local bootstrap finite sample test. In Section 6 we extend the above results to the case where the random variables of interest are multivariate. In Section 7, we propose a bootstrap bias-corrected estimator of causality measures and provide a simulation exercise to evaluate the bias-correction and investigate the finite sample properties (size and power) of local bootstrap-based test for causality measure. Section 8 is devoted to an empirical application and the conclusion relating to the results is given in Section 9. Proofs appear in the Appendix 10.

2 Motivation

The causality measures that we consider here constitute a generalization of those developed by Geweke (1982), Geweke (1984b) and others. The existing measures quantify the effect of one variable Y on another variable X assuming that the regression function linking the two variables of interest is known and is linear. These measures focus on quantifying the causality in mean and ignore the causality in high-order moments and in distribution. The significance of such measures is limited in the presence of unknown regression functions and in the presence of nonlinear and high-order moment causalities. We propose measures of Granger causality between random variables based on copula densities. Such measures detect and quantify the causalities in high-order moments and in distribution. To see the importance of such causality measures, consider the following examples.

Example 1 *Suppose we have two variables X and Y. Assume that the joint process $(X,Y)'$ follows a stationary VAR(1) model:*

$$\begin{bmatrix} X_{t+1} \\ Y_{t+1} \end{bmatrix} = \begin{bmatrix} 0.5 & 0.0 \\ 0.4 & 0.35 \end{bmatrix} \begin{bmatrix} X_t \\ Y_t \end{bmatrix} + \underbrace{\begin{bmatrix} \varepsilon^X_{t+1} \\ \varepsilon^Y_{t+1} \end{bmatrix}}_{\varepsilon_{t+1}}, \tag{1}$$

where

$$\varepsilon_{t+1} \mid X_t, Y_t \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} \sigma^2_{X,t} & 0 \\ 0 & \sigma^2_y \end{bmatrix} \right)$$

with

$$\sigma_{X,t}^2 = 0.01 + 0.5Y_t^2 + 0.25X_t^2.$$

Since the coefficient of Y_t in the first equation of (1) is zero, we can conclude that Y does not Granger cause X in the mean. However, if consider the causality in the variance we have

$$V\left(X_{t+1} \mid X_t, Y_t\right) = 0.01 + 0.5Y_t^2 + 0.25X_t^2,$$

where now Y Granger causes X in the variance. This example corresponds to the case where the causality in the variance does exist even if there is no causality in the mean. But, how can one measure the degree of the causality in the variance? Existing measures do not answer this question.

Example 2 *Suppose now X is given by the following process:*

$$X_{t+1} = \mu_X + 0.5X_t + \varepsilon_{t+1}^X,$$

where the error term ε_{t+1}^X follows Lévy skew stable probability distribution defined by the Fourier transform of its characteristic function $\varphi(u)$:

$$\varepsilon_{t+1}^X \mid X_t, Y_t \sim f\left(\varepsilon_{t+1}^X, 1, \beta_t, 1, 0\right) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \varphi(u) e^{-iu\varepsilon_{t+1}^X} du,$$

where

$$\varphi(u) = \exp\left[- \mid u \mid \left(1 + i\frac{2\beta_t \ sgn(u)}{\pi}\log(u)\right)\right],$$

and $sgn(u)$ is the sign of u and β_t is the time-varying skewness that depends on Y :

$$\beta_t = \lambda + \rho Y_t.$$

In this model, Y does not affect the mean and variance of X, but it does affect its skewness. Again, how can one measure the degree of the causality in skewness? Existing measures do not answer this question.

3 Granger causality measures

Let $\left\{(X_t, Y_t) \in \mathbb{R}\times\mathbb{R} \equiv \mathbb{R}^2,\ t = 0, ..., T\right\}$ be a sample of stationary stochastic process in $\mathbb{R}^2$, with joint distribution function F_{XY} and density function f_{XY}. For simplicity of exposition, here we consider the case of *univariate Markov processes of order one*. Later, see Section (6), we extend our results to the case where the variables of interest X and Y are *multivariate Markov processes of order p,* for $p \geq 1$.

For (X_t, Y_t) a Markov process of order one, using the results from Gouriéroux, Monfort, and Renault (1987), we define the following measure of Granger causality from X to Y:

$$C(X \to Y) = E\left\{\log\left(\frac{f\left(Y_t \mid Y_{t-1}, X_{t-1}\right)}{f\left(Y_t \mid Y_{t-1}\right)}\right)\right\}. \tag{2}$$

Important properties of this measures include: **(1)** it is nonnegative, and **(2)** it cancels only when there is no causality. The latter property means that the following null hypothesis of Granger non-causality from X to Y,

$$H_0^{x\to y} : f(y_t \mid x_{t-1}, y_{t-1}) = f(y_t \mid y_{t-1}),\ 1 \leq t \leq T, \tag{3}$$

is equivalent to $C(X \to Y) = 0$. Thus, the measure $C(X \to Y)$ defines a "distance" between left and right hand sides of the null hypothesis $H_0^{x\to y}$. Consequently, higher values of measure indicate larger causality in distribution from X to Y. Similarly, a measure of Granger causality from Y to X is defined by:

$$C(Y \to X) = E\left\{\log\left(\frac{f\left(X_t \mid X_{t-1}, Y_{t-1}\right)}{f\left(X_t \mid X_{t-1}\right)}\right)\right\}, \tag{4}$$

and the following null hypothesis of Granger non-causality from Y to X

$$H_0^{y\to x} : f(x_t \mid y_{t-1}, x_{t-1}) = f(x_t \mid x_{t-1}),\ 1 \leq t \leq T \tag{5}$$

is equivalent to $C(Y \to X) = 0$.

Finally, the instantaneous Granger causality between Y and X can also be characterized in terms of probability density functions. Thus, the instantaneous Granger *non-causality* is characterized by the following equivalent null hypotheses:

$$\begin{aligned}
H_0^{x\leftrightarrow y} &: f(y_t \mid x_t, x_{t-1}, y_{t-1}) = f(y_t \mid x_{t-1}, y_{t-1}),\ 1 \leq t \leq T,\\
H_0^{x\leftrightarrow y} &: f(x_t \mid y_t, x_{t-1}, y_{t-1}) = f(x_t \mid x_{t-1}, y_{t-1}),\ 1 \leq t \leq T,\\
H_0^{x\leftrightarrow y} &: f(y_t, x_t \mid x_{t-1}, y_{t-1}) = f(y_t \mid x_{t-1}, y_{t-1}) f(x_t \mid x_{t-1}, y_{t-1}),\ 1 \leq t \leq T.
\end{aligned} \tag{6}$$

Based on the previous null hypotheses, the instantaneous Granger causality between Y and X can be quantified using one of the following equivalent measures:

$$\begin{aligned}
C(Y \longleftrightarrow X) &= E\left\{\log\left(\tfrac{f(Y_t|X_t, X_{t-1}, Y_{t-1})}{f(Y_t|X_{t-1}, Y_{t-1})}\right)\right\}\\
&= E\left\{\log\left(\tfrac{f(X_t|Y_t, X_{t-1}, Y_{t-1})}{f(X_t|X_{t-1}, Y_{t-1})}\right)\right\}\\
&= E\left\{\log\left(\tfrac{f(Y_t, X_t|X_{t-1}, Y_{t-1})}{f(Y_t|X_{t-1}, Y_{t-1}) f(X_t|X_{t-1}, Y_{t-1})}\right)\right\}
\end{aligned} \tag{7}$$

Further, observe that:

$$C(X \to Y) + C(Y \to X) + C(Y \longleftrightarrow X) = C(Y, X) = E\left\{\log\left(\frac{f(Y_t, X_t \mid X_{t-1}, Y_{t-1})}{f\left(X_t \mid X_{t-1}\right) f\left(Y_t \mid Y_{t-1}\right)}\right)\right\}, \tag{8}$$

where the right-hand side of the equation (8) defines a measure of conditional dependence between Y and X, denoted by $C(Y, X)$. The measure of dependence $C(Y, X)$ decomposes the dependence between Y and X to measures of *feedbacks* from X to Y ($C(X \to Y)$) and from Y to X ($C(Y \to X)$) and a measure of instantaneous Granger causality between Y and X ($C(Y \longleftrightarrow X)$). It will also enable one to check whether the processes Y and X must be considered together or whether they can be treated separately.

4 Copula-based Granger causality measures

Here we show that the above causality measures can be rewritten in terms of copula densities. This will allow us to keep only the terms that involve the dependence among the random variables. Further, using copula density rather than probability density has the benefits discussed in the introduction.

It is well known from Sklar (1959) that the distribution function of any joint process $(U, V, W) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ can be expressed via a copula

$$F(u, v, w) = C\left(F_U(u), F_V(v), F_W(w)\right), \tag{9}$$

where $F_Q(.)$, for $Q = U, V, W$, is the marginal distribution function of random variable Q, and $C\left(F_U(.), F_V(.), F_W(.)\right)$ is a copula function defined on $[0,1]^3$ which captures the dependence of (U, V, W). If we derive Equation (9) with respect to (u, v, w), we obtain the density function of the joint process (U, V, W) which can be expressed as

$$f(u, v, w) = f_U(u) \times f_V(v) \times f_W(w) \times c\left(F_U(u), F_V(v), F_W(w)\right), \tag{10}$$

where $f_Q(.)$, for $Q = U, V, W$, is the marginal density of random variable Q and $c\left(F_U(.), F_V(.), F_W(.)\right)$ is a copula density defined on $[0,1]^3$ of (U, V, W). Using Equation (10), the causality measure defined in (2) can be rewritten in terms of copula densities as follows:

$$C^c(X \to Y) = E\left\{\log\left(\frac{c\left(F_{Y_t}(Y_t), F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1})\right)}{c\left(F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1})\right) c\left(F_{Y_t}(Y_t), F_{X_{t-1}}(X_{t-1})\right)}\right)\right\}, \tag{11}$$

where $c\left(F_{Y_t}(.), F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$, $c\left(F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$ and $c\left(F_{Y_t}(.), F_{X_{t-1}}(.)\right)$ are the copula densities of the joint processes (Y_t, Y_{t-1}, X_{t-1}), (Y_{t-1}, X_{t-1}), and (Y_t, X_{t-1}), respectively. We can, in a similar way, rewrite the causality measure, say $C^c(Y \to X)$, in (4) in terms of copula densities $c\left(F_{X_t}(.), F_{X_{t-1}}(.), F_{Y_{t-1}}(.)\right)$, $c\left(F_{X_{t-1}}(.), F_{Y_{t-1}}(.)\right)$ and $c\left(F_{X_t}(.), F_{Y_{t-1}}(.)\right)$.

Finally, the first measure of the instantaneous Granger causality between X and Y which is defined in (7) can be rewritten in terms of copula densities as follows:

$$C^c(X \leftrightarrow Y) = E\left\{\log\left(\frac{c\left(F_{Y_t}(Y_t), F_{X_t}(X_t), F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1})\right) c\left(F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1})\right)}{c\left(F_{X_t}(X_t), F_{X_{t-1}}(X_{t-1}), F_{Y_{t-1}}(Y_{t-1})\right) c\left(F_{Y_t}(Y_t), F_{X_{t-1}}(X_{t-1}), F_{Y_{t-1}}(Y_{t-1})\right)}\right)\right\}, \tag{12}$$

where $c\left(F_{Y_t}(Y_t), F_{X_t}(X_t), F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1})\right)$, $c\left(F_{X_t}(X_t), F_{X_{t-1}}(X_{t-1}), F_{Y_{t-1}}(Y_{t-1})\right)$ and $c\left(F_{Y_t}(Y_t), F_{X_{t-1}}(X_{t-1}), F_{Y_{t-1}}(Y_{t-1})\right)$ are the copula densities of the joint processes $(Y_t, X_t, Y_{t-1}, X_{t-1})$, (X_t, X_{t-1}, Y_{t-1}), and (Y_t, X_{t-1}, Y_{t-1}), respectively. We can, in a similar way, rewrite the last two instantaneous Granger causality measures in (7) in terms of copula densities.

5 Estimation and inference for copula-based Granger causality measures

Since we are interested in time series data, we need to specify the dependence in the processes of interest. In what follows, we consider β-mixing dependent random variables. The β-mixing condition is required to show the asymptotic normality of nonparametric estimator of Granger causality measures [see Tenreiro (1997) and Fan and Li (1999)]. However, as we will show later, the consistency of the nonparametric estimators of causality measures can be established under α-mixing condition. To prove the consistency and establish its asymptotic normality, we apply the results from Bouezmarni, Rombouts, and Taamouti (2010).

5.1 Estimation

In section 4 we have shown that the Granger causality measures can be rewritten in terms of copula densities. These measures can be estimated by replacing the unknown copula densities by their nonparametric estimates from a finite sample. Hereafter, we focus on the estimation of Granger causality measure from X to Y, $C^c(X \to Y)$, defined in (11). However, we can in a similar way define estimators of measures of Granger causality from Y to X and of the instantaneous Granger causality between X and Y defined in (12).

To estimate $C^c(X \to Y)$ in (11), we first need to estimate the copula densities $c\left(F_{Y_t}(.), F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$, $c\left(F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$ and $c\left(F_{Y_t}(.), F_{X_{t-1}}(.)\right)$. The latter can be estimated using the Bernstein copula density estimators studied by Bouezmarni, Rombouts, and Taamouti (2010) for time series and defined below. We first set the following additional notations. We denote by

$$G_t = (G_{t1}, G_{t2}, G_{t3}) = (F_{Y_t}(Y_t), F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1}))$$

and its empirical analog

$$\hat{G}_t = (\hat{G}_{t1}, \hat{G}_{t2}, \hat{G}_{t3}) = (F_{Y_t,T}(Y_t), F_{Y_{t-1},T}(Y_{t-1}), F_{X_{t-1},T}(X_{t-1})),$$

where $F_{Y_t,T}(.)$, $F_{Y_{t-1},T}(.)$, and $F_{X_{t-1},T}(.)$ with subscript T is to indicate the empirical analog of the distribution functions $F_{Y_t}(.), F_{Y_{t-1}}(.)$, and $F_{X_{t-1}}(.)$, respectively. The Bernstein copula density estimator of $c\left(F_{Y_t}(.), F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$ at a given value $g = (g_1, g_2, g_3)$ is defined by

$$\hat{c}(g_1, g_2, g_3) = \hat{c}(g) = \frac{1}{T}\sum_{t=1}^{T} K_k(g, \hat{G}_t), \tag{13}$$

where

$$K_k(g, \hat{G}_t) = k^3 \sum_{k_1=0}^{k-1}\sum_{k_2=0}^{k-1}\sum_{k_3=0}^{k-1} A_{\hat{G}_t,k} \prod_{j=1}^{3} p_{k_j}(g_j),$$

the integer k represents the bandwidth parameter, $p_{k_j}(g_j)$ is the binomial distribution

$$p_{k_j}(g_j) = \begin{pmatrix} k-1 \\ k_j \end{pmatrix} g_j^{k_j} \, (1-g_j)^{k-k_j-1}, \text{ for } k_j = 0, \cdots, k-1,$$

and $A_{\hat{G}_t,k}$ is an indicator function

$$A_{\hat{G}_t,k} = \mathbf{1}_{\{\hat{G}_t \in B_k\}}, \text{ with } B_k = \left[\frac{k_1}{k}, \frac{k_1+1}{k}\right] \times \left[\frac{k_2}{k}, \frac{k_2+1}{k}\right] \times \left[\frac{k_3}{k}, \frac{k_3+1}{k}\right].$$

The Bernstein estimators $\hat{c}\left(F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$ and $\hat{c}\left(F_{Y_t}(.), F_{X_{t-1}}(.)\right)$ of $c\left(F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$ and $c\left(F_{Y_t}(.), F_{X_{t-1}}(.)\right)$, respectively, are defined in a similar way like $\hat{c}\left(F_{Y_t}(.), F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$. Observe that the kernel $K_k(g, \hat{G}_t)$ can be rewritten as

$$K_k(g, \hat{G}_t) = \sum_{k_1=0}^{k-1} \sum_{k_2=0}^{k-1} \sum_{k_3=0}^{k-1} A_{\hat{G}_t,k} \prod_{j=1}^{3} \mathcal{B}(g_j, k_j + 1, k - k_j),$$

where $\mathcal{B}(g_j, k_j + 1, k - k_j)$ is a beta density with shape parameters $k_j + 1$ and $k - k_j$ evaluated at g_j. Observe that $K_k(g, \hat{G}_t)$ can be viewed as a smoother of the empirical density estimator by beta densities. The Bernstein copula density estimator in (13) is easy to implement, non-negative, integrates to one and is free from the boundary bias problem which often occurs with conventional nonparametric kernel estimators. Bouezmarni, Rombouts, and Taamouti (2010) establish the asymptotic bias, variance and the uniform almost convergence of Bernstein copula density estimator for α-mixing data. These properties are needed to prove the consistency and the asymptotic normality of the estimators of Granger causality measures. Notice that some other nonparametric copula density estimators are proposed in the literature. For example, Gijbels and Mielniczuk (1990) suggest nonparametric kernel methods and use the reflection method to overcome the boundary bias problem, and more recently Chen and Huang (2007) use the local linear estimator. Fermanian and Scaillet (2003) derive the asymptotic properties of kernel estimators of nonparametric copulas and their derivatives in the context of time series data.

Thus, based on the previous nonparametric estimators of copula densities, an estimator of Granger causality measure $C^c(X \to Y)$ is given by:

$$\hat{C}^c(X \to Y) = \frac{1}{T} \sum_{t=1}^{T} \left\{ \log \left(\frac{\hat{c}\left(F_{Y_t,T}(Y_t), F_{Y_{t-1},T}(Y_{t-1}), F_{X_{t-1},T}(X_{t-1})\right)}{\hat{c}\left(F_{Y_{t-1},T}(Y_{t-1}), F_{X_{t-1},T}(X_{t-1})\right) \ \hat{c}\left(F_{Y_t,T}(Y_t), F_{X_{t-1},T}(X_{t-1})\right)} \right) \right\}. \qquad (14)$$

The most basic property that the above estimator $\hat{C}^c(X \to Y)$ should have is *consistency*. To prove consistency, some regularity assumptions are needed. We consider a set of standard assumptions on the stochastic process and bandwidth parameter of the Bernstein copula density estimator.

Assumptions on the stochastic process

(A1.1) $\left\{(Y_t, X_t) \in \mathbb{R} \times \mathbb{R} \equiv \mathbb{R}^2, \ t \geq 0\right\}$, is a strictly stationary β-mixing process with coefficient $\beta_l = O(\rho^l)$, for some $0 < \rho < 1$.

(A1.2) G_t has a copula function $C\left(F_{Y_t}(.), F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$ and copula density $c\left(F_{Y_t}(.), F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$. We assume that the latter is twice continuously differentiable on $(0, 1)^3$ and bounded away from zero and bounded above.

Assumptions on the bandwidth parameter

(A1.3) We assume that for $k \to \infty$, $T\, k^{-7/2} \to 0$ and $T^{-1/2} k^{3/4} \ln(T) \to 0$.

Assumption (**A1.1**) is satisfied by many processes such as ARMA and ARCH processes as documented for example by Carrasco and Chen (2002) and Meitz and Saikkonen (2002). In Assumption **(A1.2)**, the second differentiability of $c\left(F_{Y_t}(.), F_{Y_{t-1}}(.), F_{X_{t-1}}(.)\right)$ is required by Bouezmarni, Rombouts, and Taamouti (2010) in order to calculate the bias of the Bernstein copula estimator. Assumption **(A1.3)** is needed to cancel out some bias terms and for the almost sure convergence of the Bernstein copula estimator. The bandwidth parameter k plays the inverse role compared to that of the standard nonparametric kernel, that is a large value of k reduces the bias but increases the variance. We now state the consistency of the nonparametric estimator $\hat{C}^c(X \to Y)$ in (14).

Proposition 1 *Under Assumptions **(A1.1)**-**(A1.3)**, the estimator $\hat{C}^c(X \to Y)$ defined in (14) converges in probability to $C^c(X \to Y)$.*

The proof of Proposition 1 can be found in Appendix 10. In the next section we establish the asymptotic distribution (normality) of nonparametric estimator $\hat{C}^c(X \to Y)$. This will allow us to construct tests and confidence intervals for our Granger causality measures.

5.2 Inference

Obviously, the measures proposed in the previous sections can be used to test for the Granger non-causality between random variables. Hereafter, the null hypothesis of interest is given by:

$$H_0 : C^c(X \to Y) = 0.$$

In this section, we provide the asymptotic normality of the nonparametric estimator of Granger causality measures under the above null hypothesis. We also establish the consistency of the test statistic based on the proposed estimator of Granger causality measure. Again, here we focus on the measure of Granger causality from X to Y defined in (11), but similar results can be obtained for the measures of Granger causality from Y to X and of the instantaneous Granger causality between X and Y.

Under some regular conditions [see assumptions **(A1.1)**-**(A1.3)**], the following theorem provides the asymptotic normality of nonparametric estimator $\hat{C}^c(X \to Y)$ defined in (14) under the previous null hypothesis H_0 [see the proof of Theorem 1 in Appendix 10].

Theorem 1 *Under assumptions **(A1.1)**-**(A1.3)** and H_0, we have*

$$TBE = T\, k^{-3/2} \left(2\hat{C}^c(X \to Y) - T^{-1}\, k^{3/2} \xi\right) / \sigma \xrightarrow{d} \mathcal{N}(0,1),$$

where $\sigma = \sqrt{2}\,(\pi/4)^{3/2}$ and

$$\xi = -\frac{\pi^{3/2}}{8} + \frac{\pi}{2} k^{-1/2} - k^{-1}(\pi^{1/2} - 1).$$

To prove the above Theorem 1, we follow the proof of Theorem 1 in Bouezmarni, Rombouts, and Taamouti (2011). However, it is important to notice that the bias terms, B_1, B_2 and B_3 in Theorem 1 of Bouezmarni, Rombouts, and Taamouti (2011) are estimated, whereas in the present paper these terms are calculated exactly.

For a given significance level α, we reject the null hypothesis H_0 when $TBE > z_\alpha$, where z_α is the critical value from the standard normal distribution. To test H_0, we also suggest to use the smoothed bootstrap proposed by Paparoditis and Politis (2000). The validity of a smoothed bootstrap that corresponds to a test statistic which is similar to our test statistic TBE is established in Bouezmarni, Rombouts, and Taamouti (2011) [see Proposition 3 of Bouezmarni, Rombouts, and Taamouti (2011)]. Under regular assumptions on the bootstrap kernel and the bandwidth parameter, we can show that $TBE^* \xrightarrow{d} \mathcal{N}(0,1)$, where TBE^* is the smoothed bootstrap version of TBE.

Now, we have to mention that the derivation of Theorem 1 requires the boundedness of the copula density in Assumption **A1.2**. It is true that many common families of copula are unbounded at the corners, Clayton, Gumbel, Gaussian and the Student being important examples. However, we can follow Bouezmarni, Rombouts, and Taamouti (2011) to show that the result in Theorem 1 is still valid for unbounded copula densities, if the following condition is fulfilled:

$$c(g_1, g_2, g_3) = O\left(\frac{1}{\sqrt{\prod_{j=1}^{3} g_j(1-g_j)}}\right), \tag{15}$$

where $c(g_1, g_2, g_3)$ is the copula density function that corresponds to the process (Y_t, Y_{t-1}, X_{t-1}). The condition (15) is satisfied by many common copula densities, see for example Omelka, Gijbels, and Veraverbek (2009).

Finally, the following proposition establishes the consistency of the test statistic TBE defined in Theorem 1 [see the proof of Proposition 2 in Appendix 10].

Proposition 2 *If Assumptions (**A1.1**)-(**A1.3**) hold, then the test defined in Theorem 1 is consistent for any copula $c(u,v,w)\,, c(u,v)$ and $c(u,w)$ such that:*

$$\int \log\left\{\frac{c(u,v,w)}{c(u,v)c(u,w)}\right\} dC(u,v,w) > 0,$$

where $c(u,v,w)\,, c(u,v)$ and $c(u,w)$ are the copula densities of the joint processes (Y_t, Y_{t-1}, X_{t-1}), (Y_{t-1}, X_{t-1}), and (Y_t, X_{t-1}), respectively, and $C(u,v,w)$ is the copula function of the process (Y_t, Y_{t-1}, X_{t-1}).

6 Extension: High dimensional random variables

Now let $\left\{(X_t, Y_t) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \equiv \mathbb{R}^d,\ t = 0, ..., T\right\}$ be a sample of stationary stochastic process in $\mathbb{R}^d$, where $d = d_1 + d_2$, with joint distribution function F_{XY} and density function f_{XY}. For (X_t, Y_t) a Markov process of order p, the null hypothesis of Granger non-causality from vector X to vector Y is given by

$$H_0^{X\to Y} : f(y_t \mid \underline{x}_{t-1}, \underline{y}_{t-1}) = f(y_t \mid \underline{y}_{t-1}),\ 1 \le t \le T, \tag{16}$$

where $\underline{x}_{t-1} = \{x_{t-s},\ 1 \le s \le p\}$, $\underline{y}_{t-1} = \{y_{t-s},\ 1 \le s \le p\}$, for $d_1, d_2 \ge 1$, with $y_t = (y_{1,t}, \ldots, y_{d_2,t})$ and $x_t = (x_{1,t}, \ldots, x_{d_3,t})$. Similarly, the null hypothesis of Granger non-causality from vector Y to vector X is given by

$$H_0^{Y\to X} : f(x_t \mid \underline{y}_{t-1}, \underline{x}_{t-1}) = f(x_t \mid \underline{x}_{t-1}),\ 1 \le t \le T, \tag{17}$$

where $\underset{\sim}{y}_{t-1}$ and $\underset{\sim}{x}_{t-1}$ are defined above.

The instantaneous Granger causality between the vectors Y and X can also be characterized in terms of probability density functions. Thus, the instantaneous Granger *non-causality* is characterized by the following equivalent null hypotheses:

$$
\begin{aligned}
H_0^{Y\leftrightarrow X} &: \ f(y_t \mid \underset{\sim}{x}_t, \underset{\sim}{y}_{t-1}) = f(y_t \mid \underset{\sim}{x}_{t-1}, \underset{\sim}{y}_{t-1}),\ 1 \le t \le T, \\
H_0^{Y\leftrightarrow X} &: \ f(x_t \mid \underset{\sim}{x}_{t-1}, \underset{\sim}{y}_{t}) = f(x_t \mid \underset{\sim}{x}_{t-1}, \underset{\sim}{y}_{t-1}),\ 1 \le t \le T, \\
H_0^{Y\leftrightarrow X} &: \ f(y_t, x_t \mid \underset{\sim}{x}_{t-1}, \underset{\sim}{y}_{t-1}) = f(y_t \mid \underset{\sim}{x}_{t-1}, \underset{\sim}{y}_{t-1}) f(x_t \mid \underset{\sim}{x}_{t-1}, \underset{\sim}{y}_{t-1}),\ 1 \le t \le T,
\end{aligned}
\tag{18}
$$

where now $\underset{\sim}{x}_t = \{x_{t-s},\ 0 \le s \le p\}$ and $\underset{\sim}{y}_t = \{y_{t-s},\ 0 \le s \le p\}$.

Following Section 4 and based on the previous null hypotheses (16)-(18), we have the following definitions of copula-based causality measures between random vectors X and Y.

Definition 1 *The* copula-based measure *of Granger causality from vector X to vector Y is defined as*

$$C^c(X \to Y) = E\left\{\log\left(\frac{c\left(\bar{F}_Y(Y_t), \bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1}), \bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1})\right) c\left(\bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1})\right)}{c\left(\bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1}), \bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1})\right) c\left(\bar{F}_Y(Y_t), \bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1})\right)}\right)\right\},$$

where, for simplicity of notation, we denote $\bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1}) = (F_{Y_1}(Y_{1,t-1}), ..., F_{Y_{d_2}}(Y_{d_2,t-1}), ..., F_{Y_1}(Y_{1,t-p}), ..., F_{Y_{d_2}}(Y_{d_2,t-p}))$, $\bar{F}_Y(Y_t) = (F_{Y_1}(Y_{1,t}), ..., F_{Y_{d_2}}(Y_{d_2,t}))$ *and* $\bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1}) = (F_{X_1}(X_{1,t-1}), ..., F_{X_{d_1}}(X_{d_1,t-1}), ..., F_{X_1}(X_{1,t-p}), ..., F_{X_{d_1}}(X_{d_1,t-p}))$, *with* $F_{Q_i}(.)$ *for* $Q = X, Y$, *is the marginal distribution function of the* i*-th element of the random vector* Q, *and* $c\left(\bar{F}_Y(.), \bar{F}_{\underset{\sim}{Y}}(.), \bar{F}_{\underset{\sim}{X}}(.)\right)$, $c\left(\bar{F}_{\underset{\sim}{Y}}(.), \bar{F}_{\underset{\sim}{X}}(.)\right)$, $c\left(\bar{F}_Y(.), \bar{F}_{\underset{\sim}{Y}}(.)\right)$, *and* $c\left(\bar{F}_{\underset{\sim}{Y}}(.)\right)$ *are the copula densities of the joint processes* $(Y_t,\ \underset{\sim}{Y}_{t-1},\ \underset{\sim}{X}_{t-1})$, $(\underset{\sim}{Y}_{t-1},\ \underset{\sim}{X}_{t-1})$, *and* $(Y_t,\ \underset{\sim}{Y}_{t-1})$, $\underset{\sim}{Y}_{t-1}$, *respectively.*

The *copula-based measure* of Granger causality from vector Y to vector X, say $C^c(Y \to X)$, can be defined in a similar way. We now define the measure of the instantaneous Granger causality between vectors Y and X in terms of copula densities.

Definition 2 *The* copula-based measure *of the instantaneous Granger causality between vectors Y and X is defined as*

$$C^c(X \leftrightarrow Y) = E\left\{\log\left(\frac{c\left(\bar{F}_Y(Y_t), \bar{F}_X(X_t), \bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1}), \bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1})\right) c\left(\bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1}), \bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1})\right)}{c\left(\bar{F}_Y(Y_t), \bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1}), \bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1})\right) c\left(\bar{F}_X(Y_t), \bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1}), \bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1})\right)}\right)\right\},$$

where $\bar{F}_X(X_t) = (F_{X_1}(X_{1,t}), ..., F_{X_{d_1}}(X_{d_1,t}))$, *and* $\bar{F}_Y(Y_t)$, $\bar{F}_{\underset{\sim}{Y}}(\underset{\sim}{Y}_{t-1})$, *and* $\bar{F}_{\underset{\sim}{X}}(\underset{\sim}{X}_{t-1})$ *are defined in the previous Definition (1),* $c\left(\bar{F}_Y(.), \bar{F}_X(.), \bar{F}_{\underset{\sim}{X}}(.), \bar{F}_{\underset{\sim}{Y}}(.)\right)$, $c\left(\bar{F}_Y(.), \bar{F}_{\underset{\sim}{X}}(.), \bar{F}_{\underset{\sim}{Y}}(.)\right)$, $c\left(\bar{F}_X(.), \bar{F}_{\underset{\sim}{X}}(.), \bar{F}_{\underset{\sim}{Y}}(.)\right)$, *and* $c\left(\bar{F}_{\underset{\sim}{X}}(.), \bar{F}_{\underset{\sim}{Y}}(.)\right)$ *are the copula densities of the joint processes* $(Y_t, X_t, \underset{\sim}{X}_{t-1}, \underset{\sim}{Y}_{t-1})$, $(Y_t, \underset{\sim}{X}_{t-1}, \underset{\sim}{Y}_{t-1})$, $(X_t, \underset{\sim}{X}_{t-1}, \underset{\sim}{Y}_{t-1})$ *and* $(\underset{\sim}{X}_{t-1}, \underset{\sim}{Y}_{t-1})$, *respectively.*

Finally, the *copula-based measure* of dependence between vectors Y and X can be deduced using the identity

$$C^c(X, Y) = C^c(X \to Y) + C^c(Y \to X) + C^c(Y \longleftrightarrow X).$$

Following Section (5.1) and using the Bernstein estimators of copula densities, an estimator of Granger causality measure from vector X to vector Y is given by:

$$\hat{C}^c(X \to Y) = \frac{1}{T}\sum_{t=1}^{T}\left\{\log\left(\frac{\hat{c}\left(\bar{F}_{\mathrm{Y,T}}(y_t), \bar{F}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1}), \bar{F}_{\underline{\mathrm{X}},\mathrm{T}}(\underline{\mathrm{x}}_{t-1})\right)\hat{c}\left(\bar{F}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1})\right)}{\hat{c}\left(\bar{F}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1}), \bar{F}_{\underline{\mathrm{X}},\mathrm{T}}(\underline{\mathrm{x}}_{t-1})\right)\hat{c}\left(\bar{F}_{\mathrm{Y,T}}(y_t), \bar{F}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1})\right)}\right)\right\}, \tag{19}$$

where $\bar{F}_{\mathrm{Y,T}}(y_t)$, $\bar{F}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1})$, and $\bar{F}_{\underline{\mathrm{X}},\mathrm{T}}(\underline{\mathrm{x}}_{t-1})$ with subscript T is to indicate the empirical analog of the terms $\bar{F}_{\mathrm{Y}}(y_t)$, $\bar{F}_{\underline{\mathrm{Y}}}(\underline{\mathrm{y}}_{t-1})$, and $\bar{F}_{\underline{\mathrm{X}}}(\underline{\mathrm{x}}_{t-1})$ defined in the above definitions (1)-(2). Again, here we focus on the Granger causality measure from vector X to vector Y. The estimators of measures of Granger causality from vector Y to vector X and of the instantaneous Granger causality between vectors X and Y can be defined in similar way. To prove the consistency of the estimator $\hat{C}^c(X \to Y)$ in (19), similar assumptions to the ones in section (5.1) are needed. As in Section (5), we consider a set of standard assumptions on the stochastic process and bandwidth parameter of the Bernstein copula density estimator.

Assumptions on the stochastic process

(A2.1) $\left\{(X_t, Y_t) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2} \equiv \mathbb{R}^d,\ t \geq 0\right\}$ is a strictly stationary β-mixing process with coefficient $\beta_l = O(\rho^l)$, for some $0 < \rho < 1$.

(A2.2) G_t has a copula function $C\left(\bar{F}_{\mathrm{Y}}(.), \bar{F}_{\underline{\mathrm{Y}}}(.), \bar{F}_{\underline{\mathrm{X}}}(.)\right)$ and copula density $c\left(\bar{F}_{\mathrm{Y}}(.), \bar{F}_{\underline{\mathrm{Y}}}(.), \bar{F}_{\underline{\mathrm{X}}}(.)\right)$. We assume that the latter is twice continuously differentiable on $(0,1)^{d_2+pd}$ and bounded away from zero and bounded above.

Assumptions on the bandwidth parameter

(A2.3) We assume that for $k \to \infty$, $T\, k^{-((d_2+pd)/2)-2} \to 0$ and $T^{-1/2}k^{(d_2+pd)/4}\ln(T) \to 0$.

A discussion of the above assumptions can be found in Section (5.1). We now state the consistency of the nonparametric estimator $\hat{C}^c(X \to Y)$ in (19).

Proposition 3 *Under Assumptions (**A.2.1**)-(**A.2.3**), the estimator $\hat{C}^c(X \to Y)$ defined in (19) converges almost surely to the true causality measure $C^c(X \to Y)$.*

The proof of Proposition 3 is similar to the one of Proposition 1, and thus we decided to do not include it in the paper.

Again, our null hypothesis of interest is given by:

$$H_0 : C^c(X \to Y) = 0.$$

Under the above regular conditions [see assumptions **(A2.1)-(A2.3)**], the following theorem provides the asymptotic normality of the estimator $\hat{C}^c(X \to Y)$ in (19) under the null H_0 [see the proof of Theorem 2 in Appendix 10].

Theorem 2 *Under assumptions **(A2.1)-(A2.3)** and H_0, we have*

$$TBE = T\, k^{-(d_2+pd)/2} \left(2\hat{C}^c(X \to Y) - T^{-1}\, k^{(d_2+pd)/2}\, \xi\right)/\sigma \xrightarrow{d} \mathcal{N}(0,1),$$

where $\sigma = \sqrt{2}\,(\pi/4)^{(d_2+pd)/2}$ *and*

$$\xi = -2^{-(d_2+pd)}\pi^{(d_2+pd)/2} + 2^{-(pd_2+d_2)}\pi^{(pd_2+d_2)/2}k^{-(pd_1)/2} + 2^{-pd}\pi^{pd/2}k^{-d_2/2} + 2(2^{-pd_2}-1)\pi^{pd_2/2}k^{-(pd_1+d_2)/2}.$$

The proof of the asymptotic normality in the above Theorem 2 is similar to the one of Theorem 1. Hence, in Appendix 10 we only computed the bias terms. Now, for a given significance level α, we reject the null hypothesis H_0 when $TBE > z_\alpha$, where z_α is the critical value from the standard normal distribution. To run a test and make our decision, we also use the smoothed bootstrap technique proposed by Paparoditis and Politis (2000). As mentioned before, the validity of a smoothed bootstrap that corresponds to a test statistic which is similar to our test statistic TBE is established in Bouezmarni, Rombouts, and Taamouti (2011) [see Proposition 3 of Bouezmarni, Rombouts, and Taamouti (2011)].

7 Monte-Carlo simulations

Here we examine the finite sample bias in the nonparametric estimation of Granger causality measures and we suggest a bootstrap bias-corrected estimator for these measures. We also investigate the finite sample properties (size and power) of nonparametric test proposed in Theorem 1.

7.1 Estimation: Bootstrap bias-corrected estimator of Granger causality measures

The nonparametric estimators of Granger causality measures can be biased in small samples, and this may arise from bias in copula density estimates. To correct the finite sample bias in the nonparametric estimation of Granger causality measures we suggest to use the bootstrap technique.

7.1.1 Bootstrap bias-correction

We first use bootstrap technique to compute the small sample bias in the nonparametric estimator of Granger causality measure proposed in Section 5.1. Thereafter, we subtract the bias term from the nonparametric estimate to obtain a bootstrap bias-corrected estimate of Granger causality measure. Since a simple resampling from the empirical distribution will not conserve the conditional dependence structure in the data, we suggest to use the local smoothed bootstrap proposed by Paparoditis and Politis (2000). The method is easy to implement in the following five steps:

(1) we draw the sample of $\underline{Y}^*_{t-1}$ from a nonparametric kernel (L) estimator of the density of $\underline{Y}_{t-1}$;

(2) conditional on $\underline{Y}^*_{t-1}$, we draw Y^*_t and $\underline{X}^*_{t-1}$ independently from their nonparametric kernel conditional density estimators;

(3) based on the bootstrap sample, we compute the Granger causality measure $\hat{C}^{c*}(X \to Y)$:

$$\hat{C}^{c*}(X \to Y) = \frac{1}{T}\sum_{t=1}^{T}\left\{\log\left(\frac{\hat{c}^*\left(\bar{F}^*{}_{\mathrm{Y,T}}(y_t^*), \bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1}^*), \bar{F}^*{}_{\underline{\mathrm{X}},\mathrm{T}}(\underline{\mathrm{x}}_{t-1}^*)\right)\hat{c}^*\left(\bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1}^*)\right)}{\hat{c}^*\left(\bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1}^*), \bar{F}^*{}_{\underline{\mathrm{X}},\mathrm{T}}(\underline{\mathrm{x}}_{t-1}^*)\right)\hat{c}^*\left(\bar{F}^*{}_{\mathrm{Y,T}}(y_t^*), \bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1}^*)\right)}\right)\right\},$$

where $\hat{c}^*\left(\bar{F}^*{}_{\mathrm{Y,T}}(.), \bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(.), \bar{F}^*{}_{\underline{\mathrm{X}},\mathrm{T}}(.)\right)$, $\hat{c}^*\left(\bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(.), \bar{F}^*{}_{\underline{\mathrm{X}},\mathrm{T}}(.)\right)$, $\hat{c}^*\left(\bar{F}^*{}_{\mathrm{Y,T}}(.), \bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(.)\right)$, and $\hat{c}^*\left(\bar{F}^*{}_{\underline{\mathrm{Y}},\mathrm{T}}(.)\right)$ $\left[\bar{F}^*_{\mathrm{Y,T}}(y_t^*), \bar{F}^*_{\underline{\mathrm{Y}},\mathrm{T}}(\underline{\mathrm{y}}_{t-1}^*), \text{and } \bar{F}^*_{\underline{\mathrm{X}},\mathrm{T}}(\underline{\mathrm{x}}_{t-1}^*))\right]$ are nonparametric estimators of copula densities [cumulative distributions] $c\left(\bar{F}_{\mathrm{Y}}(.), \bar{F}_{\underline{\mathrm{Y}}}(.), \bar{F}_{\underline{\mathrm{X}}}(.)\right)$, $c\left(\bar{F}_{\underline{\mathrm{Y}}}(.), \bar{F}_{\underline{\mathrm{X}}}(.)\right)$, $c\left(\bar{F}_{\mathrm{Y}}(.), \bar{F}_{\underline{\mathrm{Y}}}(.)\right)$, and $c\left(\bar{F}_{\underline{\mathrm{Y}}}(.)\right)$ $\left[\bar{F}_{\mathrm{Y}}(y_t), \bar{F}_{\underline{\mathrm{Y}}}(\underline{\mathrm{y}}_{t-1}), \text{ and } \bar{F}_{\underline{\mathrm{X}}}(\underline{\mathrm{x}}_{t-1})\right]$, respectively, computed using the bootstrap sample;

(4) we repeat the steps **(1)**-**(3)** B times so that we obtain $\hat{C}_j^{c*}(X \to Y)$, for $j = 1, ..., B$;

(5) we approximate the bias term $Bias$ $=E[\hat{C}^c(X \to Y) - C^c(X \to Y)]$ by the corresponding bootstrap bias $Bias^*$ $=E^*[\hat{C}^{c*}(X \to Y) - \hat{C}^c(X \to Y)]$, where E^* is the expectation based on the bootstrap distribution of $\hat{C}^{c*}(X \to Y)$, and $\hat{C}^c(X \to Y)$ is the estimate of $C^c(X \to Y)$ using the original sample. This suggests the bias estimate

$$\widehat{Bias}^* = \frac{1}{B}\sum_{j=1}^{B}\hat{C}_j^{c*}(X \to Y) - \hat{C}^c(X \to Y).$$

Hence, a bootstrap bias-corrected estimator of measure of Granger causality from X to Y is given by

$$\hat{C}_{BC}^{c*}(X \to Y) = \hat{C}^c(X \to Y) - \widehat{Bias}^*. \tag{20}$$

In practice and especially when the true value of causality measure is zero or close to zero, it is possible that for some bootstrap samples the quantity $\hat{C}_{BC}^{c*}(X \to Y)$ becomes negative. In this case we follow Dufour and Taamouti (2010) and suggest to impose the following non-negativity truncation:

$$\hat{C}_{BC}^{c*}(X \to Y) = \max\left\{\hat{C}_{BC}^{c*}(X \to Y), 0\right\}. \tag{21}$$

Finally, we can, in a similar way, define the bootstrap bias-corrected estimators for measures of Granger causality from Y to X and of the instantaneous Granger causality between X and Y.

7.1.2 Simulation study

We run a Monte Carlo experiment to investigate possible bias and bias-correction in the nonparametric estimation of Granger causality measures. We consider two groups of data generating processes (DGPs) that represent linear and nonlinear regression models with different forms of heteroskedasticity. Table 1 presents the DGPs used in our simulation study and Table 2 summarizes the directions of causality and non-causality in these DGPs. The first three DGPs of Y and DGP1 to DGP8 of X are used to evaluate the bias in the nonparametric estimation of Granger causality measures. In these DGPs the true values of Granger causality measures are known and equal to zero. For example, in the first three DGPs we have that Y and X are by construction independent: Y does not cause X and X does not cause Y. In DGP1 to DGP8 of X, by construction, we have Y does not cause X. Thus, we expect that Granger causality measures for these DGPs will be equal to zero. However, in DGP4 to

DGP	Y_t	X_t
DGP1	ε_{1t} white noise	ε_{2t} white noise
DGP2	$Y_t = 0.5Y_{t-1} + \varepsilon_{1t}$	$X_t = 0.5X_{t-1} + \varepsilon_{2t}$
DGP3	$Y_t = (0.01 + 0.5Y_{t-1}^2)^{0.5}\varepsilon_{1t}$	$X_t = 0.5X_{t-1} + \varepsilon_{2t}$
DGP4	$Y_t = 0.5Y_{t-1} + 0.5X_{t-1} + \varepsilon_{1t}$	$X_t = 0.5X_{t-1} + \varepsilon_{2t}$
DGP5	$Y_t = 0.5Y_{t-1} + 0.5X_{t-1}^2 + \varepsilon_{1t}$	$X_t = 0.5X_{t-1} + \varepsilon_{2t}$
DGP6	$Y_t = 0.5Y_{t-1}X_{t-1} + \varepsilon_{1t}$	$X_t = 0.5X_{t-1} + \varepsilon_{2t}$
DGP7	$Y_t = 0.5Y_{t-1} + 0.5X_{t-1}\varepsilon_{1t}$	$X_t = 0.5X_{t-1} + \varepsilon_{2t}$
DGP8	$Y_t = \sqrt{h_{1,t}}\varepsilon_{1t}$ $h_{1,t} = 0.01 + 0.5Y_{t-1}^2 + 0.25X_{t-1}^2$	$X_t = 0.5X_{t-1} + \varepsilon_{2t}$
DGP9	$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.3 \end{pmatrix} + \begin{bmatrix} 0.1 & 0.8 \\ 0.7 & 0.15 \end{bmatrix} \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$	$\begin{pmatrix} \varepsilon_t^Y \\ \varepsilon_t^X \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0.2 \\ 0.2 & 1 \end{pmatrix}\right]$

Table 1: This table summarizes the data generating processes that we consider in the simulation study to investigate the bias in the nonparametric estimation of Granger causality measures and to examine the finite sample properties (size and power) of nonparametric test for these causality measures. We simulate (Y_t, X_t), for $t = 1, \ldots, T$ under assumption that $(\varepsilon_{1t}, \varepsilon_{2t})' \sim N(0, I_2)$ and are i.i.d.

DGP9 of Y, the causality from X to Y exists. Finally, the Granger causality from Y to X also exists in GDP9 of X. Hence, in the latter DGPs the Granger causality measures from Y to X and from X to Y will not be equal to zero.

DGP	Directions of Causality
DGP1	$X \nrightarrow Y$, $Y \nrightarrow X$
DGP2	$X \nrightarrow Y$, $Y \nrightarrow X$
DGP3	$X \nrightarrow Y$, $Y \nrightarrow X$
DGP4	$X \rightarrow Y$, $Y \nrightarrow X$
DGP5	$X \rightarrow Y$, $Y \nrightarrow X$
DGP6	$X \rightarrow Y$, $Y \nrightarrow X$
DGP7	$X \rightarrow Y$, $Y \nrightarrow X$
DGP8	$X \rightarrow Y$, $Y \nrightarrow X$
DGP9	$X \rightarrow Y$, $Y \rightarrow X$

Table 2: This table summarizes the directions of causality and non-causality in the DGPs of Table 1. The symbols "$\rightarrow$" and "$\nrightarrow$" refer to Granger causality and Granger noncausality, respectively.

Obviously, the nonparametric estimators of Granger causality measures depend on the bandwidth parameter k, which is needed to estimate the copula densities. Here we take k equal to the integer part of $T^{1/2}$. In section 7.2, we consider various values of k to evaluate the sensitivity of the properties (size and power) of the tests of causality measures. To keep the computing time in our simulation study reasonable, we consider two sample sizes $T = 200$, 300. We perform 250 bootstrap replications and 500 simulations to compute the bias terms and the average values of the bootstrap bias-corrected Granger causality measures. Finally, in our simulations the data are rescaled such that the variables have zero mean and variance equal to one.

The simulation results are presented in Tables 3 and 4. From these, we see that the nonparametric estimators of Granger causality measures are biased, which is possibly due to the finite sample bias in the nonparametric estimators of Bernstein copula density. Interestingly, the results show that there is a big improvement when we use the bootstrap bias-corrected estimators.

7.2 Tests for Granger causality measure: Empirical size and power

We study the finite sample performance of the test nonparametric proposed in Theorem 1. We examine its size and power properties using the data generating processes (DGPs) introduced in section 7.1.2 [see Tables 1 and Table 2]. The first three DGPs of Y and DGP1 to DGP8 of X are used to investigate the size property, since

	DGP1	DGP2	DGP3	DGP4	DGP5	DGP6	DGP7	DGP8	DGP9
$X \to Y$	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
$\hat{C}^c(X \to Y)$	0.0633	0.0539	0.0628	0.1452	0.2337	0.1649	0.1548	0.1206	0.2986
$\hat{C}^{c*}_{BC}(X \to Y)$	0.0016	0.0000	0.0002	0.0957	0.1818	0.1024	0.0984	0.0580	0.2528
$Y \to X$	No	No	No	No	No	No	No	No	Yes
$\hat{C}^c(Y \to X)$	0.0612	0.0527	0.0539	0.0514	0.0532	0.0530	0.0548	0.0539	0.2705
$\hat{C}^{c*}_{BC}(Y \to X)$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0000	0.2253

Table 3: This table summarizes the results of bootstrap bias-corrected estimation of Granger causality measures. The sample size is equal to 200, the number of simulations is equal to 500, and the number of bootstrap replications is equal to 250. The term "No" indicates that there is non-causality in the DGP under consideration, whereas the term "Yes" means that there is causality in that GDP.

	DGP1	DGP2	DGP3	DGP4	DGP5	DGP6	DGP7	DGP8	DGP9
$X \to Y$	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes
$\hat{C}^c(X \to Y)$	0.0542	0.0467	0.0546	0.1415	0.2513	0.1638	0.1637	0.1174	0.3160
$\hat{C}^{c*}_{BC}(X \to Y)$	0.0009	0.0000	0.0003	0.0993	0.2069	0.1094	0.1150	0.0637	0.2769
$Y \to X$	No	No	No	No	No	No	No	No	Yes
$\hat{C}^c(Y \to X)$	0.0544	0.0455	0.0461	0.0448	0.0453	0.0462	0.0470	0.0466	0.2830
$\hat{C}^{c*}_{BC}(Y \to X)$	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.2447

Table 4: This table summarizes the results of bootstrap bias-corrected estimates of Granger causality measures. The sample size is equal to 300, the number of simulations is equal to 500, and the number of bootstrap replications is equal to 250. The term "No" indicates that there is non-causality in the DGP under consideration, whereas the term "Yes" means that there is causality in that GDP.

in these DGPs the null hypotheses that the Granger causality measures from X to Y and from Y to X equal to zero are satisfied. However, in DGP4 to DGP9 of Y and GDP9 of X the null hypotheses are not satisfied, and therefore serve to illustrate the power of the test.

Recall that Theorem 1 is valid only asymptotically. For finite samples and in order to improve the size and power of the proposed nonparametric test, the bootstrap technique is used to compute the test-statistics and *p-values*. As we mentioned in section 7.1.2, a simple bootstrap, *i.e.* resampling from the empirical distribution, will not conserve the conditional dependence structure in the data, and hence sampling under the null hypothesis is not guaranteed. To prevent this from occurring, we use the local smoothed bootstrap of Paparoditis and Politis (2000). The method is easy to implement in the following five steps: **(1)** we draw the sample of $\underline{Y}^*_{t-1}$ from the nonparametric kernel estimator (L) of the density of $\underline{Y}_{t-1}$; **(2)** conditional on $\underline{Y}^*_{t-1}$, we draw Y^*_t and $\underline{X}^*_{t-1}$ independently from their nonparametric kernel conditional density estimators; **(3)** using the bootstrap sample, we compute the bootstrap test-statistic TBE^* in the same way as TBE; **(4)** we repeat the steps (1)-(3) B times so that we obtain TBE^*_j, for $j = 1, ..., B$; **(5)** the bootstrap *p-value* is computed as $p^* = B^{-1}\sum_{j=1}^{B} 1_{\{TBE^*_j > TBE\}}$. For given significance level α, we reject the null hypothesis if $p^* < \alpha$.

TBE test-statistic depends on the bandwidth parameter k, which is used to estimate the copula densities. In our simulation study we take k equal to the integer part of $cT^{1/2}$, for $c = 1, 1.5, 2$. To keep the computing time in the simulations reasonable, we consider two sample sizes $T = 200$, 300 and use $B = 250$ bootstrap replications with resampling bandwidths chosen by the standard rule of thumb. Finally, we use 500 simulations to compute the empirical size and power of the test.

The size and power properties for the sample sizes 200 and 300 are given in Tables 5 and 6, respectively. For 5% and 10% significance levels and for both $T = 200$ and $T = 300$, we see that the TBE^* test controls quite well its size and has good power. For the DGP1 the test tends to be slightly oversized and is conservative for the DGP3, DGP5 and DGP6. In most cases, the power is quite good and close to 100%.

8 Empirical application: Stock market return and exchange rates

Given its crucial role in the economy, the causal relationship between stock prices and exchange rates is of great importance for academics, policymakers and professionals. In the literature, there is no academic consensus about this relationship and the results are somewhat mixed as to whether stock indexes lead exchange rates or vise versa. Most of the conclusions were obtained using *linear mean regression-based tests*. Although such tests have high power in uncovering linear causal relations, their power against nonlinear causal relations can be very low [see Hiemstra and Jones (1994), Bouezmarni, Rombouts, and Taamouti (2011) and Bouezmarni, Roy, and Taamouti (2010) among many others]. For that reason, traditional Granger causality tests might overlook a significant nonlinear relation between stock prices and exchange rates. In this section, we apply our nonparametric causality measures to quantify the causal relationship between those two variables in a broader

Table 5: Size and Power properties for sample size $T = 200$.

	DGP1	DGP2	DGP3	DGP4	DGP5	DGP6	DGP7	DGP8	DGP9
					$T = 200,\ \alpha = 5\%$				
$Y \to X$	Size	Size	Size	Size	Size	Size	Size	Size	Power
$c = 1$	5.20	6.40	5.60	5.20	5.60	4.00	4.40	4.40	100
$c = 1.5$	6.00	6.00	4.00	4.80	3.20	3.60	4.00	3.20	100
$c = 2$	6.88	4.80	2.40	3.60	4.40	4.00	4.00	4.00	100
$X \to Y$	Size	Size	Size	Power	Power	Power	Power	Power	Power
$c = 1$	6.66	4.80	5.20	100	100	100	100	99.2	100
$c = 1.5$	6.40	4.80	5.60	100	100	98.8	100	99.2	100
$c = 2$	6.95	5.20	5.60	99.2	100	98.0	100	99.6	100
					$T = 200,\ \alpha = 10\%$				
$Y \to X$	Size	Size	Size	Size	Size	Size	Size	Size	Power
$c = 1$	8.00	10.00	10.80	8.80	8.40	5.60	10.00	10.40	100
$c = 1.5$	11.20	8.40	8.40	6.00	8.40	6.00	6.80	8.80	100
$c = 2$	13.20	8.00	7.60	6.40	7.60	6.40	9.20	10.00	100
$X \to Y$	Size	Size	Size	Power	Power	Power	Power	Power	Power
$c = 1$	10.80	10.80	9.60	100	100	100	100	99.6	100
$c = 1.5$	12.00	9.60	10.00	100	100	100	100	100	100
$c = 2$	13.40	8.80	9.60	100	100	99.2	100	99.6	100

Empirical size and power at the α level based on 500 replications. The sample size is T=200 and the number of bootstrap resamples is B=250. The bandwidth k is the integer part of $cT^{1/2}$.

Table 6: Size and Power properties for sample size $T = 300$.

	DGP1	DGP2	DGP3	DGP4	DGP5	DGP6	DGP7	DGP8	DGP9
					$T = 300$, $\alpha = 5\%$				
$Y \rightarrow X$	Size	Size	Size	Size	Size	Size	Size	Size	Power
$c = 1$	6.80	3.60	2.40	5.20	4.40	4.40	3.60	5.600	100
$c = 1.5$	7.20	3.60	2.40	6.00	2.00	2.40	3.20	5.600	100
$c = 2$	5.80	3.80	2.60	5.60	3.40	2.60	4.40	4.50	100
$X \rightarrow Y$	Size	Size	Size	Power	Power	Power	Power	Power	Power
$c = 1$	6.40	6.20	6.00	100	100	100	100	100	100
$c = 1.5$	7.60	3.60	6.80	100	100	100	100	100	100
$c = 2$	5.10	4.50	5.80	100	100	100	100	100	100
					$T = 300$, $\alpha = 10\%$				
$Y \rightarrow X$	Size	Size	Size	Size	Size	Size	Size	Size	Power
$c = 1$	12.40	5.20	6.00	9.20	7.20	8.40	6.80	10.40	100
$c = 1.5$	11.60	6.80	6.40	9.60	5.20	5.60	9.20	10.40	100
$c = 2$	10.90	7.40	6.90	10.20	6.30	7.50	10.10	9.80	100
$X \rightarrow Y$	Size	Size	Size	Power	Power	Power	Power	Power	Power
$c = 1$	10.80	10.40	11.60	100	100	100	100	100	100
$c = 1.5$	11.60	8.80	11.20	100	100	100	100	100	100
$c = 2$	10.50	10.20	11.20	100	100	100	100	100	100

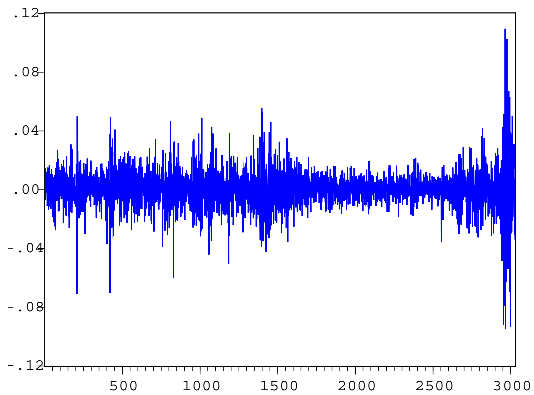
Empirical size and power at the α level based on 500 replications. The sample size is T=300 and the number of bootstrap resamples is B=250. The bandwidth k is the integer part of $cT^{1/2}$.

framework that allows us to leave free the specification of the underlying model. Thus, we compare the strength of the impact of stock market returns on growth rates of exchange rates with the strength of the impact of the latter variable on the former variable.

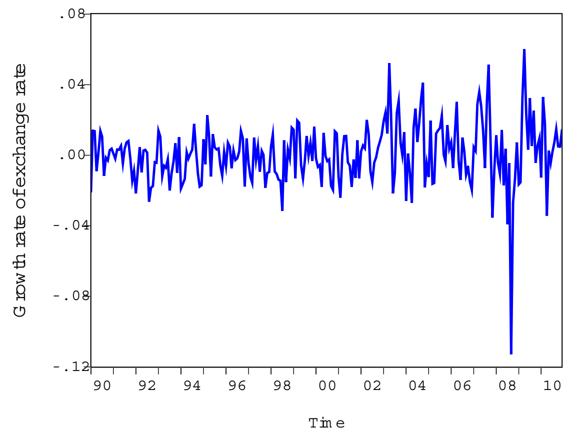
The relationship between exchange rates and stock prices have been the focus of most economic literature for quite some time. It has been recognized that exchange rate can Granger cause stock prices. To examine those causal links, early studies were using the simple correlation between the two variables. Aggarwal (1981) found that there is a positive and significant correlation between the US dollar and US stock prices, whereas Soenen and Henniga (1981) found a negative and significant correlation between the two variables. Further, Soenen and Aggarwal (1989) also found mixed results among industrial countries. Many other more recent studies have used more sophisticated econometric techniques to study stock prices-exchange rates relationships. Bahmani-Oskooee and Sohrabian (1992) using cointegration models along with Granger causality tests, found that there is bidirectional causality between stock prices measured by S&P 500 index and the effective exchange rate of the dollar, at least in the short-run. Since Bahmani-Oskooee and Sohrabian (1992) several papers have examined different directions of causality between stock prices and exchange rates using these econometric technique and data from both industrial and developing countries. The direction of causality, similar to earlier correlation studies, appears mixed. Mok (1993) using ARIMA approach and Granger causality tests, found that the Hong Kong market efficiently incorporated much of exchange rate information in its price changes both at daily market close and open. Abdalla and Murinde (1997) found out that the results for India, Korea and Pakistan suggest that exchange rates Granger cause stock prices, which is consistent with earlier study by Aggarwal (1981). But, for the Philippines, they found that the stock prices lead the exchange rates. Granger, Huang, and Yang (2000) using unit root and cointegration models found that data from South Korea are in agreement with the traditional approach. That is, exchange rates lead stock prices. On the other hand, using data of the Philippines and found that stock prices lead exchange rates with negative correlation. Further, they found that the data from Hong Kong, Malaysia, Singapore, Thailand, and Taiwan indicate strong feedback relations, whereas that of Indonesia and Japan fail to reveal any recognizable pattern. Finally, Nieh and Lee (2001), first found that there is no long-run significant relationship between stock prices and exchange rates in the G-7 countries. This result interfaces with Bahmani-Oskooee and Sohrabian (1992)'s finding, but contrasts with the studies that suggest there is a significant relationship between these two financial variables. Second, they found that the short-run significant relationship has only been found for one day in certain G-7 countries.

8.1 Data description

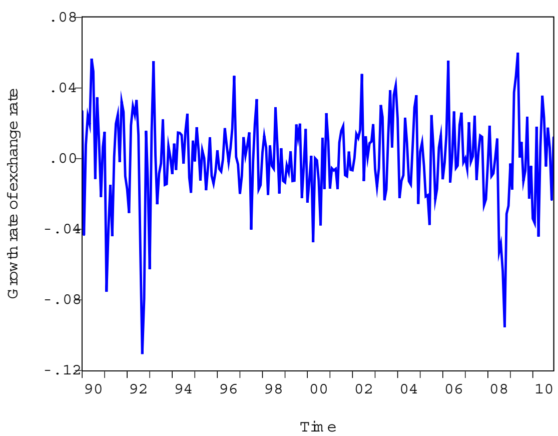
Here we use data on S&P 500 Index and US/Canada, US/UK and US/Japen exchange rates. The data sets consist of monthly observations on S&P 500 Index and exchange rates and come from St. Louis Fed and Yahoo Finance, respectively. The sample runs from January 1990 to January 2011 for a total of 253 observations, see Figure 1 for the series in growth rates.



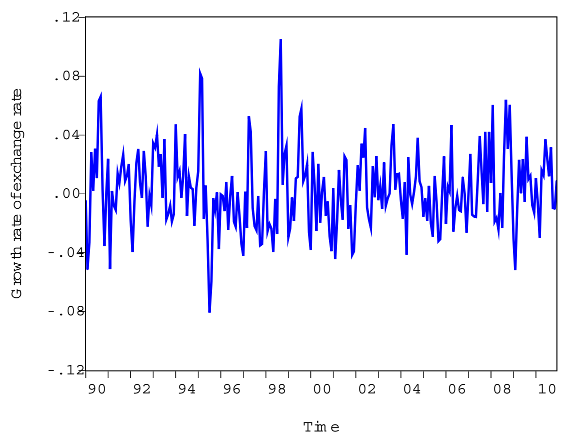
(a) Stock Return



(b) US/CA Exchange Rate



(c) US/UK Exchange Rate



(d) US/JAP Exchange Rate

Figure 1: S&P 500 stock returns and growth rates of us/canada, us/uk, and us/japan exchange rates. The sample runs from January 1990 to January 2011 for a total of 253 observations.

We perform Augmented Dickey-Fuller tests (hereafter ADF-tests) for nonstationarity of the logarithmic price and exchange rates and their first differences. Using ADF-tests with only an intercept and with both a trend and an intercept, the results show that all variables in logarithmic form are nonstationary. However, their first differences are stationary. The test statistics with both a trend and an intercept for the first differences of log price and log US/Canada, US/UK and US/Japen exchange rates are -14.666, -12.164, -11.390, -11.666, respectively, and the corresponding 5% critical value is -3.427. Using ADF-tests with only intercept leads to the same conclusions. Thus, based on the above stationarity tests we model the first difference of logarithmic price and exchange rates rather than their level. Consequently, the causality relations have to be interpreted in terms of the growth rates.

8.2 Causality measures

Here we apply the previous proposed nonparametric estimator and nonparametric test of copula-based causality measure to evaluate the strength of Granger causality from stock market return, say R, to us/i exchange rate, say EX_i for i =Canada, UK, Japan, and the strength of Granger causality from us/i exchange rate to stock market return. The estimation and inference results are reported in Tables 7 and 8. The zero-values (0.0000) of the causality measure estimates in the two tables are due to the non-negativity truncation given in 21. To compute the t-statistic we do not apply the truncation, but we use the bootstrap bias corrected estimator defined in 20.

Table 7 shows that the estimates of measures of Granger causality from $us/canada$ and us/uk exchange rates to stock market returns are *very close to zero*. Consequently, the effects of $us/canada$ and us/uk exchange rates on stock market returns are *economically very weak*. These effects are statistically *insignificant* at 5% significance level. Table 7 also shows that the impact of $us/japan$ exchange rate on stock market returns is *economically strong*, and it is statistically *significant* at 5% significance level. Now from Table 8 we see that the impacts of stock market returns on us/uk and $us/japan$ exchange rates are *economically weak* and statistically *insignificant* at 5% significance level. Further, we see that the effect of stock market returns on $us/canada$ exchange rate is *economically strong* and statistically significant at 5% significance level. Finally, the comparison of the previous two tables shows that the impact of stock market market returns on exchange rates is *much stronger* than the impact of exchange rate on stock market returns.

9 Conclusion

We proposed a nonparametric estimator and a nonparametric test for Granger causality measures that quantify linear and nonlinear causality in distribution between random variables. We first showed that the Granger causality measures can be rewritten in terms of copula densities. Thereafter, we proposed consistent nonparametric estimators for these Granger causality measures based on consistent nonparametric estimators of copula

Table 7: Measures of Causality from Exchange Rates to Stock Returns

	Bias corrected estimate	T-test
$C^c(EX_{Canada} \to R)$		
$c = 1$	0.0019	0.2006
$c = 1.5$	0.0000	-0.0832
$c = 2$	0.0000	-0.5919
$C^c(EX_{UK} \to R)$		
$c = 1$	0.0000	-0.9407
$c = 1.5$	0.0000	-0.9313
$c = 2$	0.0000	-0.5820
$C^c(EX_{Japan} \to R)$		
$c = 1$	0.0293	**3.0216**
$c = 1.5$	0.0345	**2.7683**
$c = 2$	0.0347	**2.2594**

Note: This table reports the results of the bootstrap bias-corrected estimation and T-test of measures of Granger causality from exchange rates to stock market returns .The zero-values (0.0000) of the causality measure estimates are due to the non-negativity truncation given in 21. To compute the t-statistic we do not apply the truncation, but we use the bootstrap bias corrected estimator defined in 20.

Table 8: Measures of Causality from Stock Returns to Exchange Rates

	Bias corrected estimate	T-test
$C^c(R \to EX_{Canada})$		
$c = 1$	0.0357	**3.6097**
$c = 1.5$	0.0330	**2.6113**
$c = 2$	0.0284	**1.9763**
$C^c(R \to EX_{UK})$		
$c = 1$	0.0105	1.0497
$c = 1.5$	0.0095	0.7370
$c = 2$	0.0097	0.6314
$C^c(R \to EX_{Japan})$		
$c = 1$	0.0083	0.85171
$c = 1.5$	0.0071	0.57494
$c = 2$	0.0082	0.54371

Note: This table reports the results of the bootstrap bias-corrected estimation and T-test of measures of Granger causality from stock market returns to exchange rates. The zero-values (0.0000) of the causality measure estimates are due to the non-negativity truncation given in 21. To compute the t-statistic we do not apply the truncation, but we use the bootstrap bias corrected estimator defined in 20.

densities. We proved that the nonparametric estimators of the measures are asymptotically normally distributed and we discussed the validity of a local smoothed bootstrap that can be used in finite sample settings to compute bootstrap bias-corrected estimators and build tests for Granger causality measures. A simulation study revealed that the bootstrap bias-corrected estimator of causality measures behaves well and that the test has quite good finite sample size and power properties for a variety of typical data generating processes and different sample sizes. Finally, we illustrated the practical relevance of nonparametric causality measures by quantifying the Granger causality between S&P500 Index returns and many exchange rates: US/Canada, US/UK and US/Japen exchange rates.

10 Appendix: Proofs

This Appendix provides the proofs of the theoretical results developed in sections 5-6. Except for the proof of Proposition 1, most of the rest of the proofs here are inspired from the paper Bouezmarni, Rombouts, and Taamouti (2011).

Proof of Proposition 1

First, we denote $\xi_t = \big(F_{Y_t}(y_t), F_{Y_{t-1}}(y_{t-1}), F_{X_{t-1}}(x_{t-1})\big)$ and $\xi_{t,T} = \big(F_{Y_t,T}(y_t), F_{Y_{t-1},T}(y_{t-1}), F_{X_{t-1},T}(x_{t-1})\big)$. Using Taylor expansion and the fact that $|\xi_{t,T} - \xi_t| = O_P(T^{-1/2})$ uniformly, we obtain

$$\log(\hat{c}(\xi_{t,T})) = \log(\hat{c}(\xi_t)) + O_P(T^{-1/2}). \tag{22}$$

Second, using Taylor expansion and the fact that $|\hat{c}(\xi_t) - c(\xi_t)| = O_P(k^{-1} + T^{-1/2}k^{3/4}\ln(T))$ uniformly [see Bouezmarni, Rombouts, and Taamouti (2010)], we have

$$\log(\hat{c}(\xi_t)) = \log(\hat{c}(\xi_t)) + O_P(k^{-1} + T^{-1/2}k^{3/4}\ln(T)). \tag{23}$$

Thus, from (22) and (23), we obtain

$$\hat{C}^c(X \to Y) = \frac{1}{T}\sum_{t=1}^{T}\left\{\log\left(\frac{c\left(F_{Y_t}(Y_t), F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1})\right)}{c\left(F_{Y_{t-1}}(Y_{t-1}), F_{X_{t-1}}(X_{t-1})\right) c\left(F_{Y_t}(Y_t), F_{X_{t-1}}(X_{t-1})\right)}\right)\right\} + O_P(\eta(k,T)),$$

where $\eta(k,T) = T^{-1/2} + k^{-1} + T^{-1/2}k^{3/4}\ln(T)$. Hence, the results of Proposition 1 can be deduced from the law of large numbers.

■

Proof of Theorem 1:

Without loss of generality and since

$$\hat{c}(\bar{F}_{Y,T}(y), \bar{F}_{\underline{Y},T}(\underline{y}), \bar{F}_{\underline{X},T}(\underline{x})) = \hat{c}(\bar{F}_Y(y), \bar{F}_{\underline{Y}}(\underline{y}), \bar{F}_{\underline{X}}(\underline{x})) + O_P(T^{-1}),$$

where the convergence $O_P(T^{-1})$ is uniform in $(0,1)^d$ and $\bar{F}_\zeta(.)$ ($\bar{F}_{\zeta,T}(.)$) is the distribution function of ζ (resp. the empirical distribution function of ζ) with ζ is either Y, $\underline{Y}$ or $\underline{X}$. In what follows, we will work with

$$H(\hat{c}, \hat{C}) = \frac{1}{T}\sum_{t=1}^{T}\log\left\{\frac{\hat{c}(G_t)}{\hat{c}(U_t)\hat{c}(V_t)}\right\},$$

where

$G_t = (\bar{F}_Y(Y_t), \bar{F}_{\underline{\mathrm{Y}}_t}(\underline{\mathrm{Y}}_{t-1}), \bar{F}_{\underline{\mathrm{X}}}(\underline{\mathrm{X}}_{t-1}))$, $U_t = (\bar{F}_Y(Y_t), \bar{F}_{\underline{\mathrm{Y}}}(\underline{\mathrm{Y}}_{t-1}))$, and $V_t = (\bar{F}_{\underline{\mathrm{Y}}}(\underline{\mathrm{Y}}_{t-1}), \bar{F}_{\underline{\mathrm{X}}}(\underline{\mathrm{X}}_{t-1}))$.

First, we show the asymptotic normality of $H(\hat{c}, C) \equiv \int \log\left\{\frac{\hat{c}(u,v,w)}{\hat{c}(u,v)\hat{c}(u,w)}\right\} dC(u,v,w)$. Second, we state that $(H(\hat{c}, \hat{C}) - H(\hat{c}, C))$ is negligible. We start with the following lemma

Let us consider We start with Taylor expansion, we obtain

$$\begin{aligned} H(\hat{c}, C) &\approx \int \left(\frac{\hat{c}(u,v,w)}{\hat{c}(u,v)\hat{c}(u,w)} - 1\right) dC(u,v,w) - \frac{1}{2}\int \left(\frac{\hat{c}(u,v,w)}{\hat{c}(u,v)\hat{c}(u,w)} - 1\right)^2 dC(u,v,w) \\ &+ \frac{1}{6}\int \left(\frac{\hat{c}(u,v,w)}{\hat{c}(u,v)\hat{c}(u,w)} - 1\right)^3 dC(u,v,w) \\ &= I_1 + I_2 + I_3. \end{aligned} \tag{24}$$

We denote by

$$\phi(\alpha) = \frac{\phi_3(\alpha)}{\phi_1(\alpha)\phi_2(\alpha)} - 1$$

with

$$\phi_1(\alpha) = c(u,v) + \alpha c^*(u,v), \phi_2(\alpha) = c(u,w) + \alpha c^*(u,w), \text{ and } \phi_3(\alpha) = c(u,v,w) + \alpha c^*(u,v,w)$$

where $c^*(u,v,w)$, $c^*(u,v)$ and $c^*(u,w)$ are functions in Γ_i for $i = 1$, 2 and 3 respectively and Γ_i is a set defined as

$$\Gamma_i = \left\{\gamma : [0,1]^{q_i} \to \mathbb{R},\ \gamma \text{ is bounded},\ \int \gamma = 0\right\}$$

Using Taylor's expansion, we have

$$\phi(\alpha) = \phi(0) + \alpha\phi'(0) + \frac{1}{2}\alpha^2\phi''(\alpha^*), \text{ for } \alpha^* \in [0, \alpha].$$

We can show that:

$$\phi'(\alpha) = \frac{c^*(u,v,w)\phi_1(\alpha)\phi_2(\alpha) - c^*(u,v)\phi_2(\alpha)\phi_3(\alpha) - c^*(u,w)\phi_1(\alpha)\phi_3(\alpha)}{\phi_1^2(\alpha)\phi_2^2(\alpha)},$$

and

$$\phi''(\alpha) = O(||c^*(u,v)c^*(u,v,w)||_\infty + ||c^*(u,w)c^*(u,v,w)||_\infty).$$

Next, we consider $\alpha = 1$, $c^*(u,v,w) = \hat{c}(u,v,w) - c(u,v,w)$, $c^*(u,v) = \hat{c}(u,v) - c(u,v)$, and $c^*(u,w) = \hat{c}(u,w) - c(u,w)$. Using the results of Bouezmarni, Rombouts, and Taamouti (2010), we get

$$\phi''(\alpha) = O_p\left(\zeta(k,T)\right)$$

with $\zeta(k,T) = T^{-1}k^{5/4}\ln^2(T)$. Under H_0, we have $\phi(0) = 0$ and

$$\phi'(0) = \frac{\hat{c}^*(u,v,w)}{c(u,v,w)} - \frac{\hat{c}(u,v)}{c(u,v)} - \frac{\hat{c}(u,w)}{c(u,w)} + 1.$$

Hence, we have

$$
\begin{aligned}
I_1 &= \int \left(\frac{\hat{c}(u,v,w)}{\hat{c}(u,v)\hat{c}(u,w)} - 1 \right) dC(u,v,w) \\
&= \int \left(\frac{\hat{c}(u,v,w)}{c(u,v,w)} - \frac{\hat{c}(u,v)}{c(u,v)} - \frac{\hat{c}(u,w)}{c(u,w)} + 1 \right) dC(u,v,w) + O_p\left(\zeta(k,T)\right) \\
&= 2 - \int \frac{\hat{c}(u,v)}{c(u,v)} c(u,v,w) dudvdw - \int \frac{\hat{c}(u,w)}{c(u,w)} c(u,v,w) dudvdw + O_p\left(\zeta(k,T)\right) \\
&= 2 - \int \frac{\hat{c}(u,v)}{c(u,v)} c(u,v) dudv - \int \frac{\hat{c}(u,w)}{c(u,w)} c(u,w) dudw + O_p\left(\zeta(k,T)\right) \\
&= O_p\left(\zeta(k,T)\right).
\end{aligned}
$$

Therefore,

$$I_1 = o_p(Tk^{-3/2}).$$

Also, we can show that

$$I_3 = o_p(Tk^{-3/2}).$$

Hence, the asymptotic distribution of $H(\hat{c}, C)$ is given by that of I_2. From Bouezmarni, Rombouts, and Taamouti (2011),

$$\frac{T\,k^{-3/2}}{\sigma}\left(2I_2 - C_1 T^{-1} k^{3/2} - B_1 T^{-1} k - B_2 T^{-1} k B_3 T^{-1} k^{1/2}\right) \to \mathcal{N}(0,1), \tag{25}$$

where, $C_1 = -2^{-3}\pi^{3/2}$, and $\sigma = \sqrt{2}\,(\pi/4)^{3/2}$, $B_1 = B_2 = \frac{\pi}{4}$ and $B_3 = 1 - \pi^{1/2}k^{1/2}$.

■

Proof of Proposition 2: The proof of Proposition 2 can be deduced from the proof of Theorem 1 by observing that the term

$$\phi(0) = \frac{c(u,v,w)}{c(u,v)c(u,w)} - 1 \approx log\left(\frac{c(u,v,w)}{c(u,v)c(u,w)}\right)$$

and by the assumption $log\left(\frac{c(u,v,w)}{c(u,v)c(u,w)}\right) > 0$, we obtain that the test statistics, TBE, converges to infinity.

■

Proof of Theorem 2

The proof of the asymptotic normality for the high dimensional random variables is similar to the one in Theorem 1, where $\zeta(k,T) = T^{-1}k^{(2d_1+2d_2+d_3)/4}\ln^2(T) + T^{-1}k^{(2d_1+d_2+2d_3)/4}\ln^2(T)$ and equation (25) is given by

$$\frac{T\,k^{-l/2}}{\sigma}\left(2I_2 - C_1 T^{-1} k^{l/2} - B_1 T^{-1} k^{(l_1+l_2)/2} - B_2 T^{-1} k^{(l_1+l_3)/2} - B_3 T^{-1} k^{l_1/2}\right) \xrightarrow{d} \mathcal{N}(0,1), \tag{26}$$

with $l_1 = pd_2, l_2 = d_2, l_3 = pd_1$ and $l = l_1 + l_2 + l_3$. Here, we only focus on the calculation of the bias terms. Observe that , $C_1 = -2^{-l}\pi^{l/2}$, and $\sigma = \sqrt{2}\,(\pi/4)^{l/2}$ in (26). For the bias terms B_1 and B_2, we obtain, by denoting

$\mathbf{u} = (u_1, ..., u_{l_1}), \mathbf{v} = (v_1, ..., v_{l_2}), \mathbf{w} = (w_1, ..., w_{l_3})$

$$\begin{aligned}
B_1 &= 2^{-(l_1+l_2-1)}\pi^{(l_1+l_2)/2} + \int \frac{\prod_{j=1}^{l_1}\prod_{i=1}^{l_2}(4\pi\, u_j(1-u_j)v_i(1-v_i))^{-1/2}}{c(\mathbf{u},\mathbf{v})} c(\mathbf{u},\mathbf{v},\mathbf{w})\, d\mathbf{u}d\mathbf{v}d\mathbf{w} + o(k^{(l_1+l_2)/2}),\\
&= 2^{-(l_1+l_2-1)}\pi^{(l_1+l_2)/2} + \int \frac{\prod_{j=1}^{l_1}\prod_{i=1}^{l_2}(4\pi\, u_j(1-u_j)v_i(1-v_i))^{-1/2}}{c(\mathbf{u},\mathbf{v})} c(\mathbf{u},\mathbf{v})\, d\mathbf{u}\, d\mathbf{v} + o(k^{(l_1+l_2)/2}),\\
&= 2^{-(l_1+l_2-1)}\pi^{(l_1+l_2)/2} + \int \prod_{j=1}^{l_1}\prod_{i=1}^{l_2}(4\pi\, u_j(1-u_j)v_i(1-v_i))^{-1/2}\ d\mathbf{u}\, d\mathbf{v} + o(k^{(l_1+l_2)/2}),\\
&= 2^{-(l_1+l_2)}\pi^{(l_1+l_2)/2}.
\end{aligned}$$

and

$$B_2 = 2^{-(l_1+l_3)}\pi^{(l_1+l_3)/2}.$$

For the last term B_3 and from the proof of Lemma 5 in Bouezmarni, Rombouts, and Taamouti (2011) we have $B_3 = -(D_4 + D_7 + D_8 + D_9 + D_{10})$, with

$$\begin{aligned}
D_4 &= k^{l_1/2}\int \frac{\prod_{j=1}^{l_1}(4\pi\, u_j(1-u_j))^{-1/2}}{c(u)} c(\mathbf{u},\mathbf{v},\mathbf{w})\, d\mathbf{u}d\mathbf{v}d\mathbf{w} + o(k^{l_1/2})\\
&= k^{l_1/2}\int \prod_{j=1}^{l_1}(4\pi\, u_j(1-u_j))^{-1/2}\ d\mathbf{u} + o(k^{l_1/2})\\
&= 2^{-l_1}\pi^{l_1/2}k^{l_1/2},
\end{aligned}$$

$$D_7 = D_8 = 2\,2^{-l_1}\pi^{l_1/2}k^{l_1/2},$$

$$\begin{aligned}
D_9 &= -2k^{l_1/2}\int \frac{c(v)c(s)}{c(\mathbf{u},\mathbf{v})}(4\pi)^{-l_1/2}\prod_{j=1}^{l_1}(u_j(1-u_j))^{-1/2} d\mathbf{w}d\mathbf{v}d\mathbf{u} + o(k^{l_1/2})\\
&= -2k^{l_1/2}\int (4\pi)^{-l_1/2}\prod_{j=1}^{l_1}(u_j(1-u_j))^{-1/2} d\mathbf{u} + o(k^{l_1/2})\\
&= -2\,2^{-l_1}\pi^{l_1/2}k^{l_1/2},
\end{aligned}$$

and

$$\begin{aligned}
D_{10} &= -2k^{l_1/2}\int \frac{c(w)c(s)}{c(u,w)}(4\pi)^{-l_1/2}\prod_{j=1}^{l_1}(u_j(1-u_j))^{-1/2} d\mathbf{v}d\mathbf{w}d\mathbf{u} + o(k^{l_1/2})\\
&= -2\,2^{-l_1}\pi^{l_1/2}k^{l_1/2},
\end{aligned}$$

Therefore, $B_3 = -(2 - 2^{-l_1+1})\pi^{l_1/2}k^{l_1/2}$. This concludes the proof of the theorem.

■

References

ABDALLA, I., AND V. MURINDE (1997): "Exchange Rate and Stock Price Interactions in Emerging Financial Markets: Evidence on India, Korea, Pakistan, and Philippines," *Applied Financial Economics*, 7, 25–35.

Aggarwal, R. (1981): "Exchange rates and stock prices: A study of the US capital markets under floating exchange rates," *Akron Business and Economic Review*, 12, 7–12.

Bahmani-Oskooee, M., and A. Sohrabian (1992): "Stock Prices and the Effective Exchange Rate of the Dollar," *Applied Economics*, 24, 459–464.

Boudjellaba, H., J.-M. Dufour, and R. Roy (1992): "Testing Causality Between Two Vectors in Multivariate ARMA Models," *Journal of the American Statistical Association*, 87, 1082–1090.

——— (1994): "Simplified Conditions for Non-Causality Between Two Vectors in Multivariate ARMA Models," *Journal of Econometrics*, 63, 271–287.

Bouezmarni, T., J. Rombouts, and A. Taamouti (2010): "Asymptotic Properties of the Bernstein Density Copula Estimator for Alpha-Mixing Data," *Journal of Multivariate Analysis*, 101, 1–10.

——— (2011): "A Nonparametric Copula Based Test for Conditional Independence with Applications to Granger Causality," *forthcoming in Journal of Business & Economic Statistics*.

Bouezmarni, T., R. Roy, and A. Taamouti (2010): "Nonparametric Tests for Conditional Independence Using Conditional Distribution," *Universidad Carlos III de Madrid and University of Montreal Working Paper*.

Carrasco, M., and X. Chen (2002): "Mixing and Moment Properties of Various GARCH and Stochastic Volatility Models," *Econometric Theory*, 18, 17–39.

Chen, S. X., and T. Huang (2007): "Nonparametric Estimation of Copula Functions for Dependent Modeling," *Canadian Journal of Statistics*, 35, 265–282.

Dufour, J.-M., and A. Taamouti (2010): "Short and Long Run Causality Measures: Theory and Inference," *Journal of Econometrics*, 154, 42–58.

Fan, Y., and Q. Li (1999): "Central Limit Theorem for Degenerate U-Statistics of Absolutely Regular Processes with Applications to Model Specification Tests," *Journal of Nonparametric Statistics*, 10, 245–271.

Fermanian, J., and O. Scaillet (2003): "Nonparametric Estimation of Copulas for Time Series," *Journal of Risk*, 5, 25–54.

Gabaix, X., P. Gopikrishnan, H. V. Plerou, and E. Stanley (2003): "A Theory of Power Law Distributions in Financial Market Fluctuations," *Nature*, 423, 267–70.

Geweke, J. (1982): "Measurement of Linear Dependence and Feedback Between Multiple Time Series," *Journal of the American Statistical Association*, 77, 304–313.

——— (1984a): "Inference and Causality in Economic Time Series," *in Z. Griliches and M. D. Intriligator, eds, Handbook of Econometrics, North-Holland, Amsterdam*, 2, 1102–1144.

——— (1984b): "Measures of Conditional Linear Dependence and Feedback Between Time Series," *Journal of the American Statistical Association*, 79, 907–915.

Gijbels, I., and J. Mielniczuk (1990): "Estimating The Density of a Copula Function," *Communications in Statistics - Theory and Methods*, 19, 445–464.

Gouriéroux, C., and A. Monfort (1997): "Time Series and Dynamic Models," *Cambridge University Press, Cambridge, U.K.*

Gouriéroux, C., A. Monfort, and É. Renault (1987): "Kullback Causality Measures," *Annales dÉconomie et de Statistique*, 6/7, 369–410.

Granger, C., B. Huang, and C. Yang (2000): "A Bivariate Causality between Stock Prices and Exchange Rates: Evidence from Recent Asian Flu," *The Quarterly Review of Economics and Finance*, 40, 337354.

Granger, C. W. J. (1969): "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–459.

Hiemstra, C., and J. Jones (1994): "Testing for Linear and Nonlinear Granger Causality In The Stock Price-Volume Relation," *Journal of Finance*, 49, 1639–1664.

Lütkepohl, H. (1991): "Introduction to Multiple Time Series Analysis," *Springer-Verlag, Berlin.*

McCloskey, D. N., and S. Ziliak (1996): "The Standard Error of Regressions," *The Journal of Economic Literature*, XXXIV, 97–114.

Meitz, M., and P. Saikkonen (2002): "Ergodicity, Mixing, and Existence of Moments of a Class of Markov Models with Applications to GARCH and ACD Models," *Econometric Theory*, 24, 1291–1320.

Mok, H. (1993): "Causality of interest rate, exchange rate and stock prices at stock market open and close in Hong Kong," *Asia Pacific Journal of Management*, 10, 123–143.

Newbold, P. (1982): "Causality Testing in Economics," *in O. D. Anderson, ed., Time Series Analysis: Theory and Practice 1, North-Holland, Amsterdam.*

Nieh, C., and C. Lee (2001): "Dynamic relationship between stock prices and exchange rates for G7 countries," *Quarterly Review of Economics and Finance*, 41, 477490.

Omelka, M., I. Gijbels, and N. Veraverbek (2009): "Improved Kernel Estimation of Copulas: Weak Convergence and Goodnees-of-Fit Testing," *Annals of Statistics*, 37, 3023–3058.

Paparoditis, E., and D. Politis (2000): "The Local Bootstrap for Kernel Estimators under General Dependence Conditions," *Annals of the Institute of Statistical Mathematics*, 52, 139–159.

Pierce, D. A., and L. Haugh (1977): "Causality in Temporal Systems: Characterizations and Survey," *Journal of Econometrics*, 5, 265–293.

Polasek, W. (1994): "Temporal Causality Measures Based on AIC," *in H. Bozdogan, ed., Proceedings of the Frontier of Statistical Modeling: An Informal Approach, Kluwer, Dordrecht, The Netherlands*, pp. 159–168.

——— (2002): "Bayesian Causality Meaures for Multiple ARCH Models Using Marginal Likelihoods," *Technical report, Institute of Statistics and Econometrics, Univeristy of Basel., Switzerland.*

Sancetta, A., and S. Satchell (2004): "The Bernstein Copula and its Applications to Modeling and Approximating of Multivariate Distributions," *Econometric Theory*, 20, 535–562.

Sklar, A. (1959): "Fonction de Répartition à n Dimensions et leurs Marges," *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229–231.

Soenen, L., and R. Aggarwal (1989): "Financial Prices as Determinants of Changes in Currency Values," *Paper presented at the 25th Annual Meetings of Eastern Finance Association, Philadelphia.*

Soenen, L., and r. E. Henniga (1981): "An Analysis of Exchange Rates and Stock Prices: The US Experience Between 1980 and 1986," *Akron Business and Economic Review*, 19, 71–76.

Tenreiro, C. (1997): "Loi Asymptotique Des Erreurs Quadratiques Intégrées Des Estimateurs a Noyau de la Densite et de la Régression Sous Des Conditions de Dépendance," *Portugaliae Mathematica*, 54, 197–213.

Wiener, N. (1956): "The Theory of Prediction," *in E. F. Beckenback, ed., The Theory of Prediction, McGraw-Hill, New York, chapter 8.*