Physica A 398 (2014) 280–288

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

A permutation entropy based test for causality: The volume–stock price relation

Mariano Matilla-García[a,*], Manuel Ruiz Marín[b], Mohammed I. Dore[c]

[a] Paseo Senda del Rey, 11. Department of Quantitative Methods. UNED, 28040, Spain

[b] Universidad Politécnica de Cartagena, Spain

[c] Brock University, Canada

HIGHLIGHTS

- A novel statistic for detecting causality in nonlinear processes is presented.
- The statistic captures causality in conditional mean and in conditional variance.
- The new test is compared with other well-established tests.
- The causal relation between trade volume and stock prices is studied.

ARTICLE INFO

Article history:
Received 26 July 2013
Received in revised form 11 November 2013
Available online 8 December 2013

Keywords:
Causality test
Permutation entropy
Symbolic dynamics

ABSTRACT

The purpose of this paper is to propose a newly developed non-parametric test for linear and nonlinear causality based on permutation entropy and to show its usefulness in analyzing the potential causal relationship between trading volume and security prices. Most of the empirical applications and tests for causality rely on using Granger causality based test for linear models. Although these tests have high power in uncovering linear causal relations, their power against nonlinear causal relations can be low. Our test is designed to deal with the detection of linear and non-linear causality. We also compare our permutation entropy based test with other Granger causality tests. Monte Carlo simulations show excellent performance (in terms of size and power) of the new test for detecting linear and non-linear causality under different scenarios. Our conclusions point that there is a bidirectional causal relation from volume to price returns not only in the mean but also in the variance.

1. Introduction

The analysis of trading volume and its relationship with security prices and changes in price is a topic that has been considered for over 50 years. Its roots are generally credited to the work of Osborne [1], who initiated a long line of work that considered the possible relationship between returns and the volume of trading. Volume is a measure of the quantity of shares that change owners for a given security. The amount of daily volume on a security can fluctuate on any given day depending on the amount of new information available about the company [2]. Abnormally large volumes can be due to differences in the investor's view of the valuation after incorporating the new information. The analysis of trading volume and associated price changes corresponding to informational releases has been of much interest to researchers in prices of equities, bonds, or other derivative securities.

* Corresponding author. Tel.: +34 913987215; fax: +34 913986335.
E-mail address: mmatilla@cee.uned.es (M. Matilla-García).

0378-4371/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.physa.2013.11.031

There are several well established explanations for the presence of a causal relation between stock prices and trading volume. (i) It provides insight into the information structure of financial markets; (ii) it is important for event studies that use a combination of price and volume data from which to draw inferences; (iii) it is critical to the debate over the empirical distribution of speculative prices; and (iv) it has several significant implications for research into futures prices. (i)–(iv) are documented in Ref. [3]. In Ref. [4] it is shown that several financial models have been accordingly developed with these four points and many empirical studies supported that relation between trading volume and change price. Obviously, statistical tests for causality can provide useful information on the nature of the price–volume dynamic relationship. Although there are some papers that explicitly test for causality between prices and volume (see Ref. [4]), the vast majority has to face that problem of using a traditional linear Granger-based tool for a financial world that is almost certainly nonlinear. Among the notable exceptions that consider the nonlinear nature of the financial market are [5,6].

In general, Granger causality based tests can only give information about linear features of the stochastic processes, and thus higher order causality structures are unfortunately neglected (see Ref. [7] and references there in for a recent critical survey). Extensions to nonlinear cases now exist, but these extensions can be more difficult to use in practice and their statistical properties are less well understood. For example, in Ref. [8] the globally nonlinear data is divided into a locally linear neighborhood, whereas Ref. [9] used a radial basis function method to perform a global nonlinear regression. One factor limiting a nonlinear extension of Granger causality tests is that some economic series (for example, financial series) display a broad spectrum of characteristics that makes it difficult to produce a general framework of nonlinear models capable of capturing these characteristics. This observation suggests that any potential model can be misspecified, which is another pitfall that affects both linear and nonlinear approaches. An interesting contribution is [4], which develops nonlinear Granger causality test based on correlation integrals.

In this paper we introduce a nonparametric causality test that allows one to test for linear and nonlinear causal relationships and, given its generality, we use it for testing causality in daily volume and returns data for the Standard and Poor's 500 index (SP500). To this end we rely on and extend the approach given in Ref. [10] that uses permutation entropy to test for independence between time series. Permutation entropy is a particular type of symbolic analysis [11], which has very interesting properties for the study of dynamical systems [12] and it has been used previously in Refs. [13,14,5,15,16], within in a causality framework, although in a different way as we do in this paper. Like in those papers, we rely on symbolic analysis, that uses an information-theoretic measure, namely, permutation entropy, which is based on Shannon's entropy. In particular, transforming observed series into a finite set of symbols avoids theoretical problems that arise when dealing with continuous random variables and stochastic processes. Particularly, as in Ref. [16], our method can be understood as a type of conditional mutual information which also uses ordinal patterns. In contrast with other treatments of causality, we aim to provide a tool (a statistical test) that can be used for stochastic models that have causal dependence either in the conditional mean or in the conditional variance, and the researcher does not know where the potential causality might come from. This concern is motived because real financial time series are usually studied taking into account both conditional moments. To implement the new method it is necessary to rely on bootstrapping methods, as it is aimed to remain in a free-model approach and also to reduce the number of assumptions at minimum level. These characteristics make a distinction with other mentioned available tools that also used ordinal patterns for analyzing causality. From this point of view, the paper is a further step in the understanding of the statistical properties and utilities of permutation entropy, as it is the case of [16] and the issue where it is published.

The paper is structured as follows. The next section is an introduction to the notation and to symbolic dynamics, including permutation entropy. In Section 3, the causality test statistic is developed. In Section 4, the process of determining the optimal lag structure of the data generating process is outlined. In Section 5, some Monte Carlo experiments of different data generating processes (DGPs) are presented and compared with the classical linear Granger causality test and with Hiemstra and Jones (HJ) for nonlinear Granger causality. Section 6 is devoted to the application of our test to testing for causality in SP500 index and daily traded volume. The final section contains the conclusions.

2. Preliminaries

In this section we follow the same notation as in Ref. [10], which is based on the seminal notion of permutation entropy [11]. Let $\{X_t\}_{t\in I}$ be a real-valued time series. For a positive integer $m \geq 2$ we denote by S_m the symmetric group of order $m!$, that is the group formed by all the permutations of length m. Let $\pi_i = (i_1, i_2, \ldots, i_m) \in S_m$. We will call an element π_i in the symmetric group S_m a symbol. The positive integer m is usually known as *embedding dimension*.

Now we define an ordinal pattern for a symbol $\pi_i = (i_1, i_2, \ldots, i_m) \in S_m$ at a given time $t \in I$ such that $t + i \in I$ for all $i = 1, 2, \ldots, m - 1$. To this end we consider that the time series is embedded in an m-dimensional space as follows:

$$X_m(t) = (X_{t+0}, X_{t+1}, \ldots, X_{t+m-1}) \quad \text{for } t \in I.$$

Then we say that t is of π_i-type if and only if $\pi_i = (i_1, i_2, \ldots, i_m)$ is the unique symbol in the group S_m satisfying the two following conditions:

(a) $X_{t+i_1} \leq X_{t+i_2} \leq \cdots \leq X_{t+i_m}$, and
(b) $i_{s-1} < i_s$ if $X_{t+i_{s-1}} = X_{t+i_s}$.