

An Independence Test Based on Joint Recurrences

Teresa Aparicio, Eduardo F. Pozo*, Dulce Saura

Departamento de Análisis Económico, Universidad de Zaragoza, Zaragoza, Spain
Email: epozo@unizar.es

Received 24 July 2015; accepted 16 August 2015; published 19 August 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We propose in this paper a test procedure to determine whether two series proceed from independent systems or not. Our starting point is a multivariate extension of the methodology called Recurrence Quantification Analysis (RQA). We derive the test procedure from the probability distribution of the number of joint recurrences of both series under the null hypothesis of independence. The behavior of the test is evaluated by means of a large set of simulations, carried out with different types of dynamical systems: random, deterministic chaotic, deterministic non-chaotic, systems affected by noise and coupled systems. We obtain satisfactory results in all cases. Finally, the methodology is used to study two questions, on which the bulk of the existing economic literature agrees: 1) the relationship between the nominal interest rate and the inflation rate; and 2) the relationship between the gross domestic product and the employment. The results suggest that our test can be a suitable tool for detecting linear and nonlinear dependence between real series.

Keywords

Independence Test, Joint Recurrence Plot, Recurrence Quantification Analysis, Dynamical Systems

1. Introduction

In Economics, when working with time series we understand that they are the result of an unknown generating process, and alternative methods to ascertain the properties of the underlying dynamics are necessary. This is particularly relevant when it comes to analyzing the possible interrelation between two (or more) variables, which is a primary research topic in Economics. Different procedures have been proposed in the literature to test

*Corresponding author.

the interrelations or statistical dependencies among time series: these include the widely used Pearson, Spearman, and Kendall correlation coefficients, Hoeffding D-measure, methods based on mutual information, and a good number of more recent contributions: distance correlation [1], the maximal information coefficient [2], the Heller-Heller-Gorfine measure [3], Copula-based methods¹, and more.

The best known is the Pearson correlation coefficient, which is also simple to obtain and interpret. Nevertheless, this tool only serves to detect linear dependencies. The Spearman and the Kendall rank correlation coefficients work well in identifying linear and monotonic nonlinear relationships, but they are less than adequate in detecting other kinds of nonlinearities. Besides, if the data set is large, assigning ranks to long series is costly in calculation terms, and using rank correlation measures is problematic when the series have a large number of repeated values (which is common with the value zero in many financial series). In this case, the results being obtained can be deceptive. The other methods being mentioned are valid, in principle, for the detection of all types of relationships, although some are also computationally complex and others, like those based on Copula functions, are also difficult to implement. All these approaches are valid only in asymptotic terms, since the exact probability distributions of the corresponding test statistics are unknown. A comparative study of some of these methods is found in [5].

In this work, we propose a method to test for the existence of statistical dependence (linear or nonlinear) between two time series whose data-generating processes are unknown. We wish to emphasize that, as we will show later, our test procedure presents two advantages: its simplicity of implementation and interpretation, and its basis in an exact probability distribution.

Our starting point is the methodology based on the so-called Recurrence Plots, RP, (Eckmann *et al.* 1987) [6], and more specifically on the Recurrence Quantification Analysis, RQA, (Zbilut and Webber, 1992 [7] and Webber and Zbilut 1994 [8]). Both RP and RQA have been used to analyze the properties of dynamical systems in Physics, Chemistry, Biology, and, more recently, in Economics. Some references from the field of Economics are [9]-[18]. A review of the applications of this technique to different fields can be found in [19], where its increasing use in Economics is noted, particularly in the study of Financial Markets.

The analysis of the recurrences provides very useful information, as much about the structure of the underlying systems as about the possible links or interrelations among them. This last is studied through the multivariate extensions of the RQA: the so-called “cross recurrences” and “joint recurrences”. Our test procedure is specifically based on these latter.

The rest of the paper is structured as follows. In Section 2, we give a brief explanation of the methodology used. In Section 3, we choose a measure of the joint recurrences and obtain its probability distribution under the hypothesis of independence. In Section 4, a statistical test using this distribution is defined. Sections 5 and 6 deal with the application of the proposed method to simulated and economic series respectively. Finally, Section 7 provides a summary of our main conclusions.

2. The Recurrence Plot and the Recurrence Quantification Analysis

As we have indicated in the previous section, the starting point of this work is the so-called Recurrence Plot, a graphic tool that displays the recurrences of a data series (x), constructed using the Recurrence Matrix, RM:

$$RM_{i,j} = \Theta\left(r - \|x_i - x_j\|\right), \quad i, j = 1, \dots, n, \quad (1)$$

where Θ is the Heaviside function, r is a predefined distance threshold (or “radius”), and x_i, x_j are observations, or system states. Clearly, the Recurrence Matrix, RM , is a symmetric $n \times n$ matrix², with the values “1” of this matrix being the RP points. The aim of this tool is to detect patterns of recurrence in the data, one of the most important characteristics of dynamical systems.

In order to improve the quality of the information provided by the RPs, Zbilut and Webber (1992) [7] and Webber and Zbilut (1994) [8] quantify the information underlying the RP by means of a series of indicators based on the point density and on the structures that form certain lines in the plot. This methodology is called “Recurrence Quantification Analysis” (RQA). More recently, two extensions of this same methodology have

¹Although the theoretical fundamentals of this methodology are found in Sklar (1959) [4], their application to the field of economics (specifically to finances) has not been relevant until recent years.

²It would also be possible to use a variable radius, different for each x_i . In this case, the recurrence matrix would not be symmetrical. In [20], different procedures to determine the radius are proposed.

been proposed to analyze the relationship between different systems. The first of these is the Cross Recurrence Plot (CRP), designed by Zbilut *et al.* (1998) [21] to analyze dependencies between two systems by comparing their states. To do this, the Cross Recurrence Matrix (*CRM*) is used, defined as:

$$CRM_{i,j} = \Theta\left(r - \|x_i - y_j\|\right), \quad i = 1, \dots, n_1; \quad j = 1, \dots, n_2, \quad (2)$$

where x_i are the observations from the first system and y_j are the observations from the second system. The CRP is obtained from this matrix and shows the recurrences of each of the states of one of the systems with each of the states of the other.

The second extension is the Joint Recurrence Plot (JRP), proposed by Romano (2004) [22] to analyze the interrelation between two systems. The starting point in this case is the individual recurrences of the observations of each system, from which the Joint Recurrence Matrix (*JRM*) is constructed, defined as:

$$JRM_{i,j} = \Theta\left(r_x - \|x_i - x_j\|\right) \cdot \Theta\left(r_y - \|y_i - y_j\|\right), \quad i, j = 1, \dots, n \quad (3)$$

The *JRM* is a symmetric matrix, since both RMs are symmetric. The values “1” of this matrix correspond to the simultaneous recurrences of both systems, hence being the points of the JRP. So, if both systems are related in some way, they must present a high number of joint recurrences. On the contrary, if they are not, or if the relationship is very weak, the number of joint recurrences should be small.

Marwan *et al.* (2007) [23] study and compare the CRP and the JRP, concluding that “CRPs are more appropriate to investigating relationships between the parts of the same system... On the other hand, JRPs are more appropriate for the investigation of two interacting systems that influence each other”.

Our main goal in this article is to obtain a criterion that enables us to distinguish whether two dynamical systems are independent or not. To do this, we use a measure obtained from the joint recurrences, the “Joint Percentage of Recurrence” or “Joint Recurrence Rate”, *JREC*, which summarizes the most relevant information of the JRP.

This same approach is adopted in two recent works, [24] and [25], in which the joint recurrences are used to test interrelations and connectivity between real series. The authors introduce an indicator that they denominate “Recurrence-Based Measure of Dependence”, from which they construct a simple statistical test to determine whether both series are (or are not) connected. This statistical test is carried out using surrogate data sets, so its probability distribution is not exact, having only asymptotic validity. Our test procedure is based on the same methodology, it is equally easy to apply and interpret, and it is based on an exact probability distribution, which implies that it can be used for data series of any length.

3. The *JREC* Statistic and Its Probability Distribution

As mentioned above, the JRP shows the joint recurrences of two series, x and y . Given that its construction requires that both recurrence matrices have the same order, the series that we will use to build the JRP must be of equal size, n . Since individual recurrence matrices are symmetric and all points on their main diagonal are trivially recurrent, the same will occur with the *JRM* matrix. Thus, from now on, only the upper (or lower) triangular part of these matrices, excluding the main diagonal, will be used. In the rest of the paper, when we refer to *RM* and *JRM*, it should be understood that we are dealing with their corresponding triangular parts.

The main indicator that summarizes the information contained in the JRP is the *JREC*, defined as $JREC = 100 (NJR/NP)$, where *NJR* is the number of recurrence points of the JRP, or number of 1s of the *JRM*, and *NP* is the total of the elements of the *JRM*. Obviously, $NP = n(n - 1)/2$.

We denote by NR_x and NR_y the number of recurrence points of each series (x and y), and REC_x and REC_y their corresponding recurrence rates. Hereafter, if individual recurrence ratios are different, subindex 1 will be assigned to the RP with the smaller recurrence, and subindex 2 to the RP with the larger recurrence. It is then evident that $0 \leq JREC \leq REC_1 \leq REC_2$.

The greater the *JREC* value, the larger the percentage of simultaneous recurrences. Consequently, in general terms, if both series have been generated from systems with a high degree of dynamical similarity, the *JREC* value will be close to the maximum $(REC_1)^3$. By contrast, if both series have been generated from independent

³According to Romano (2004) [22] and Marwan *et al.* (2007) [23], it could be said that the systems are “synchronized”.

systems, the *JREC* value would be small. Although in principle one could think that, in this case, *JREC* should be equal to zero, this is not necessarily true. Even in series coming from completely independent systems, joint recurrences can appear by chance.

However, in practice, we do not have a proximity criterion. Additionally, for intermediate values, this indicator would not lead us to conclusive results. Consequently, a statistical test providing a cut-off point that allows us to establish whether both series are independent or not, would be necessary.

In order to construct this test, we will obtain the probability distribution of the random variable *NJR* for a JRP built from two series generated in independent processes (for example, two purely random series), which will give a structure and a number of points of the JRP that are absolutely random. This assumption of independence will be the null hypothesis of the test.

Under the null, the probability that a point of the first recurrence plot, RP_1 , is also a point of the second one, RP_2 , is NR_2/NP . And, obviously, the probability that a point of the RP_1 is not a point of the RP_2 is $(NP-NR_2)/NP$. Consequently, the probability that no point of the RP_1 is a point of the RP_2 , and hence $P(NJR = 0)$, is given by the expression:

$$P(NJR = 0) = \frac{\prod_{i=0}^{(NR_1-1)} (NP - NR_2 - i)}{\prod_{i=0}^{(NR_1-1)} (NP - i)} \quad (4)$$

The probability that only one point of the RP_1 is also a point of the RP_2 , *i.e.* $P(NJR = 1)$, is given by the product of three terms:

$$P(NJR = 1) = \frac{NR_2}{NP} \cdot \frac{\prod_{i=0}^{(NR_1-2)} (NP - NR_2 - i)}{\prod_{i=1}^{(NR_1-1)} (NP - i)} \cdot NR_1$$

The first term is the probability that a point of the RP_1 is also a point of the RP_2 . The second term is the probability that the remaining points of the RP_1 are not points of the RP_2 . The result is multiplied by NR_1 , the number of different cases that can appear. After rearranging, we obtain:

$$P(NJR = 1) = \frac{\prod_{i=0}^{(NR_1-2)} (NP - NR_2 - i)}{\prod_{i=0}^{(NR_1-1)} (NP - i)} \cdot NR_2 \cdot NR_1$$

Reasoning in a similar way, the probability that there are two points in the JRP would be:

$$P(NJR = 2) = \left(\frac{NR_2}{NP} \cdot \frac{NR_2 - 1}{NP - 1} \right) \cdot \frac{\prod_{i=0}^{(NR_1-3)} (NP - NR_2 - i)}{\prod_{i=2}^{(NR_1-1)} (NP - i)} \cdot \binom{NR_1}{2}$$

The first term is the probability that two points of the RP_1 are also points of the RP_2 . The second term is the probability that the remaining points of the RP_1 are not points of the RP_2 . The third term is the number of possible different combinations of pairs of recurrent points of the RP_1 . Rearranging terms, we obtain:

$$P(NJR = 2) = \frac{\prod_{i=0}^{(NR_1-3)} (NP - NR_2 - i)}{\prod_{i=0}^{(NR_1-1)} (NP - i)} \cdot NR_2 \cdot (NR_2 - 1) \cdot \binom{NR_1}{2}$$

In general terms for $NJR = k$ ($0 < k < NR_1$) we would obtain the following expression:

$$P(NJR = k) = \prod_{i=0}^{k-1} (NR_2 - i) \cdot \frac{\prod_{i=0}^{(NR_1-k-1)} (NP - NR_2 - i)}{\binom{NR_1-1}{(NR_1-1)}} \cdot \binom{NR_1}{k}, \quad (5)$$

for $k = 1, 2, 3, \dots, NR_1 - 1$.

Finally, for $NJR = NR_1$ the resulting probability would be:

$$P(NJR = NR_1) = \frac{\prod_{i=0}^{(NR_1-1)} (NR_2 - i)}{\binom{NR_1-1}{(NR_1-1)}} \cdot \prod_{i=0}^{(NR_1-1)} (NP - i). \quad (6)$$

This probability distribution is the key tool that will enable us to test whether, in a JRP obtained from two real series, the number of points differs significantly from the expected if both series had been generated from independent processes (the null hypothesis of our test). Hence, rejecting the null hypothesis would be evidence, for the chosen significance level, that both series are related, presenting a certain degree of statistical dependence or dynamical similarity.

4. Testing Procedure

Our test procedure is based on a comparison of the *JREC* value of the chosen series with the expected value of the *JREC*, $E(JREC)$, of two series generated from independent systems. $E(JREC)$ can be trivially obtained from $E(NJR)$, *i.e.* the mathematical expectation of the probability distribution derived in the previous section. Note that, even under independence, points in the JRP may appear (“false non-independences”). In general terms, the number of these false non-independences will be greater, the greater the number of recurrence points in the individual RPs. Hence, the existence of a high number of points in the JRP does not necessarily mean statistical dependence. This would only be accepted if the number of points of the JRP is significantly greater than the number we would expect if both series were independent. In this case, we would conclude that there is evidence to reject the null hypothesis, that is to say, to accept that there exists a relationship (of some kind) between the series.

In order to apply our test procedure, the following five steps must be performed:

1) Select the two-series object of study, which can be simulated or real. In previous analyses, we have verified that the existence of trend (deterministic or stochastic) in the series may give rise to trivial recurrences. To avoid these “spurious” recurrences the real series should be detrended.

2) Construct the RPs of both series. In order to do this, it is necessary to set the value of the radius r . This radius is a threshold, or distance, that determines the criterion of proximity. Schinkel *et al.* (2008) [20] state that, “for JRPs the threshold should be chosen in a fashion that the recurrence point density is the same in the individual RPs”. As in our previous work ([14], [15] and [16]), we choose a value of r to achieve a portion of recurrence points equal to 10%. Thus, in all cases, $REC_1 = REC_2 = 10^4$.

3) Construct the JRP and compute the number of recurrence points, *NJR*.

4) Obtain the probability distribution of *NJR*, using expressions (4), (5) and (6). As it has already been demonstrated, the only required data are NP , NR_1 and NR_2 .

5) Calculate the p -value of the test, *i.e.* the lowest significance level for which the null hypothesis should be rejected.

To illustrate the procedure described above, we will apply it initially to the two following cases:

First, the series x and y obtained from a simulation of the Henon chaotic system⁵; as both were generated in the same system, we expect the result of our test to show a clear relationship.

Second, two series obtained respectively from the Henon and the Ikeda chaotic systems; these series come from two different systems, so the results of the test should reflect their independence.

The results obtained, for $n = 200, 500$, and 1000 , are shown in the following table.

⁴Nevertheless, in numerous experiments, we have confirmed that our results are very robust to the choice of the radius value.

⁵The equations of the dynamical systems used in this work are included in **Appendix**.

Table 1. Results of the application of the test to two pairs of simulated series.

Series	n	NJR	$E(NJR)$	$JREC$	$E(JREC)$	p -value
Henon x - Henon y	200	764	199	3.8391	1	0.0000
	500	4962	1247	3.9775	1	0.0000
	1000	19,904	4995	3.9848	1	0.0000
Henon - Ikeda	200	178	199	0.89	1	0.9564
	500	1148	1247	0.92	1	0.9993
	1000	4735	4995	0.95	1	1

Table 1 shows:

- The number of points (NJR), and the percentage of recurrence ($JREC$) of the JRP;
- The expected average values of NJR and $JREC$ if both systems were independent, $E(NJR)$ and $E(JREC)$;
- The p -values associated with each of the tests.

The results obtained are conclusive:

For the two series generated from the Henon system, the NJR and $JREC$ values are approximately four times greater than would have been expected if both series were independent. The p -values are equal to zero, so that the null hypothesis is clearly rejected. This occurs in all three sample sizes. For the series generated from different systems, Henon and Ikeda, the values of the indicators are very close to those expected under the null hypothesis. The large p -values obtained in this case provide no evidence against the null hypothesis for the three sample sizes, at any level of significance.

These results can be considered an illustration of the good behavior of the test procedure in two extreme cases. Next, in order to better study its performance, we will use the same systems transformed in the following way:

First, we apply the test procedure to the series x and y of the Henon system, affected in this case with different levels of additive noise (as a percentage of the standard deviation of the series). As is obvious, noise distorts the structure of the system, so the higher the level of noise, the weaker should be the relationship between the series. This would have to be reflected in the results of our test.

Second, we apply the test procedure to the unidirectionally coupled Henon and Ikeda systems. Specifically, we consider the Ikeda system to be driven by the Henon system, with different levels of coupling. Each value of the coupling strength parameter reflects the amount of coupling between the systems. Our test should reflect the fact that the greater the value of this strength parameter, the greater will be the NJR and $JREC$ values, and thus the greater the interdependence between both systems.

The results obtained, with a sample size $n = 500$ are shown in Table 2:

The results displayed in Table 2 show that, in the case of the Henon system, if the noise level increases, the joint recurrences decrease. As could be expected, the interdependence between both series is maintained for small levels of noise, and diminishes as the noise level increases, until it is completely lost when the series are strongly distorted by noise. This last happens for a level of noise close to 80%.

For the Henon - Ikeda coupled system, the results also agree with what we could expect: if the strength parameter increases, so does the interdependence between both systems. For small values of the strength parameter, both systems behave as independent (the null hypothesis could not be rejected). This is the situation until this parameter reaches a value of about 0.1, when we fail to reject the null hypothesis at a significance level of 0.05. From this parameter value onwards, the level of dependence between the systems increases continuously, reaching, for greater values of the parameter, p -values equal to zero. In these cases, when the null hypothesis cannot be rejected at any level of significance, we can state that our test concludes that both systems are totally coupled.

Obviously, only a few systems and one simulation in each instance are not sufficient to evaluate the adequacy of the testing procedure. Thus, more systems and a greater number of replicas should be used. This will be done in the next section.

5. Application of the Test Procedure to Simulated Series

In order to analyze the reliability of the proposed method, the previous examples should be extended in two ways:

Table 2. Results of the application of the test to the series of the Henon system affected by noise, and to a pair of series of the Henon - Ikeda coupled system [$E(NJR) = 1247$, $E(JREC) = 1$].

% Noise	Henon x - Henon y			Strength	Henon - Ikeda		
	NJR	JREC	p -value		NJR	JREC	p -value
0	4853	3.89	0.0000	0	1148	0.92	0.9993
0.01	4819	3.86	0.0000	0.001	1190	0.95	0.9666
0.1	3307	2.65	0.0000	0.01	1270	1.02	0.2439
0.2	2370	1.90	0.0000	0.1	1295	1.04	0.0702
0.5	1539	1.23	0.0000	0.5	1370	1.10	0.0001
0.8	1292	1.04	0.0836	1	1374	1.10	0.0000
0.9	1263	1.01	0.3176	2	1401	1.12	0.0000
1	1183	0.0095	0.9801	3	1406	1.13	0.0000

- 1) Use a wide set of systems that permit more situations of both interrelation and independence to be covered;
- 2) Repeat each experiment many times.

In this section, the test procedure is applied to simulated series from the following systems:

- Two chaotic deterministic systems (Henon and Ikeda);
- One non-chaotic deterministic system with strange attractor (Gopy)⁶;
- Two periodic deterministic systems (Sine and Cosine);
- Two random systems (Normal and Uniform).

Different combinations of these systems allow us to obtain pairs of series on which a priori independence could be expected, and other pairs in which a relationship would exist.

In the simulations arising from the systems of Henon, Ikeda, and Gopy (the three bi-dimensional) one of the two variables has been selected, except in those denoted “Henon x - Henon y ” and “Ikeda x -Ikeda y ”, in which we apply the test procedure to two series coming from the same system. In one of the experiments, we attempt to determine whether our test is capable of detecting the relationship existing between series of the same system, but with different embedding dimension⁷. In the experiment we denote as “Linear transformation”, a series obtained from the distribution $N(0, 1)$ and a linear combination from this same series is used. Finally, as in the previous Section, we will use the series x and y of the Henon system affected by noise, and the series generated from the Henon - Ikeda coupled system.

Table 3 and **Table 4** summarize the results obtained for the set of simulations, presenting the percentages of rejection of the null hypothesis of independence at the 5% significance level, and the average p -value.

Table 3 has been structured in two groups. The first corresponds to six pairs of systems in which a priori independence could be expected, while the second contains another six in which a relationship clearly exists.

The results obtained allow us to establish the following conclusions:

1) For the first group, the percentage of rejection is very small in all cases, always less than 8%. This indicates a good performance; when the null hypothesis of independence is true, it is rejected only a very small percentage of times. The average p -values are high in all cases, being in the range that would lead us to fail to reject the null hypothesis for any level of significance. The results are quite similar for the three sample sizes considered (the larger sample sizes improve the results of the test in some cases, but not all).

2) For the second group, the percentage of rejection is 100%, except in one case, in which it is 97.2%. In other words, the null hypothesis (which in this case is false) is rejected in practically all cases. Clearly, the average p -values obtained lead us to the same conclusion. These results provide clear evidence of interrelation between the pairs of series of this second group.

Table 4 shows, first, that our test is capable of detecting that the dependence between the series generated by the same system decreases when the noise level increases. Specifically, the result of the test shows that this

⁶See [26].

⁷That is, applying the Time Delay Method proposed by Packard *et al.* (1980) [27]. This method transforms the original series into a series formed by m -tuples, m being the so-called embedding dimension. As is shown in Takens (1981) [28], under certain conditions, both systems would be topologically equivalent.

Table 3. Percentages of rejection of the null hypothesis of independence ($\alpha = 5\%$) and average p -values for the systems selected. Number of replicas = 500.

	$n = 200$		$n = 500$		$n = 1000$	
	% reject.	Avg. p -value	% reject.	Avg. p -value	% reject.	Avg. p -value
Sine - Henon	0.8	0.9141	0.2	0.9430	0.2	0.9262
Henon - Ikeda	7.4	0.5361	6.9	0.5581	0.4	0.9510
Ikeda - GOPY	5.8	0.5637	4.6	0.5990	0	0.9938
Normal - Uniform	4.2	0.5234	3.4	0.5167	4.8	0.6882
Normal - Sine	6.8	0.5194	2.2	0.7875	0	0.9847
Uniform - Henon	5.2	0.5029	6.4	0.5010	6.6	0.5115
Linear Transformation	100	0.0000	100	0.0000	100	0.0000
Henon 1 - 3 embed. dim.	100	0.0000	100	0.0000	100	0.0000
Sine-Cosine	100	0.0000	100	0.0000	100	0.0000
Sine-Sine squared	100	0.0000	100	0.0000	100	0.0000
Henon x - Henon y	100	0.0000	100	0.0000	100	0.0000
Ikeda x -Ikeda y	97.2	0.0059	100	0.0000	100	0.0000

Table 4. Percentages of rejection of the null hypothesis of independence ($\alpha = 5\%$) and average p -values for the series of the Henon system affected by noise and the series of the Henon - Ikeda coupled system. Number of replicas = 500.

Henon x - Henon y			Henon - Ikeda		
% Noise	% reject.	Avg. p -value	Strength	% reject.	Avg. p -value
0	100	0	0	6.9	0.5581
0.01	100	0	0.001	13.0	0.4763
0.1	100	0	0.01	96.6	0.0241
0.2	100	0	0.1	100	0
0.5	100	0	0.5	100	0
0.8	81	0.0515	1	100	0
0.9	63.6	0.1366	2	100	0
1	33.8	0.3718	3	100	0
1.5	8.2	0.5653	4	100	0
2	6.4	0.5913	5	100	0

relationship is maintained until the level of noise is high (more than 80%). Second, for the coupled system, when the strength parameter increases, the percentage of rejections of the null hypothesis also increases. It is observed that, even for small levels of coupling, the existence of a relationship is detected in almost 100% of the cases.

These results suggest that our test procedure is robust in all studied situations. This good performance for simulated series indicates that it can be a useful tool for detecting possible interrelations or dependence between real series. This is the objective of the following section, in which we illustrate the application of our test procedure to some examples with economic series.

6. Application of the Test Procedure to Economic Series

In this Section, we use this methodology to analyze the possible interdependence between certain economic se-

ries. Specifically, we focus on two issues that have been dealt with through different approaches in the economic literature, on which there is consensus:

- 1) The relationship between inflation rate and nominal interest rates;
- 2) The relationship between employment and GDP.

We emphasize that we do not intend to study either of these questions in depth, which would require a broader and more thorough examination; our objective here is only to exemplify some of the possibilities and utilities offered by our test procedure.

6.1. The Relationship between Nominal Interest Rates and Inflation Rate

This relationship is widely studied in the literature, and has been dealt with from multiple points of view, for a broad range of variables. Regarding interest rates, different approaches have been used: nominal and real, ex-post and ex-ante, different maturity terms, diverse time frequencies... Similarly, the inflation rate can be the actual or the expected rate. In the case of considering the expected variables, questions associated with the formation of expectations also appear.

From among all the cases studied in the empirical literature, there is consensus on the existence of a positive relationship between the nominal interest rate and the actual inflation rate.

We now apply the methodology described in this work to monthly series of the US economy, to check whether or not this relationship is detected.

Specifically, we use the following series:

- 1) Nominal interest rates:

—Secondary market rate of the 3-month Certificate of Deposit, chosen as an example of a short-term interest rate. Period: January 1934 to March 2015.

—One year Treasury Constant Maturity Rate, a medium-term interest rate. Period: April 1953 to March 2015.

—Moody's Seasoned AAA Corporate Bond Yield, representative of long-term interest rates. Period: January 1919 to March 2015⁸.

- 2) Inflation rate: rate of the interannual variation of the CPI-U (Consumer Price Index for All Urban Consumers) published by the US Bureau of Labor Statistics.

Table 5 presents our results, clearly confirming the existence of a relationship between the three series of interest rates and the inflation rate for the chosen periods. The three *p*-values are zero, which means that the null hypothesis should be rejected for any level of significance.

We could consider whether this strong relationship would be maintained for certain sample sub-periods: a first period up to 1953, which includes the Great Depression and the two World Wars; a second period from 1953 to 1980; and a third period from 1980 to 2015. We consider that this third sub-period corresponds to a greater development of the financial markets, the appearance of new instruments, an increase in the level of information accessible to agents, and a high mobility of capital among countries.

The obtained results are shown in **Table 6**:

We observe that the existence of a relationship between the two variables is maintained for all the sub-periods considered. Nevertheless, we see a noticeable variability in the percentage of recurrence among the different series, as well as among the different sub-periods. In all cases, this indicator is somewhat higher in the second sub-period than in the other two. Additionally, if we discount the first sub-period, from which there is only information for one of the series, in general we notice that the relationship is greater for long-term interest rates. It is generally assumed that this is the rate that best responds to market forces, so it is reasonable that the relationship to the inflation rate is stronger. In any case, to study the causes of these results in depth, which we understand

Table 5. Application of the test to series of nominal interest rates against the series of the inflation rate.

Series	<i>n</i>	<i>NREC</i>	<i>E(NREC)</i>	<i>REC</i>	<i>E(REC)</i>	<i>p</i> -value
Short term rate—Inflation rate	975	6707	4748	1.41	1	0.0000
Medium term rate—Inflation rate	744	4831	2764	1.75	1	0.0000
Long term rate—Inflation rate	1155	11,122	6664	1.67	1	0.0000

⁸The three interest rates series are available on the Economic Research section of the Federal Reserve Bank of St Louis's website.

Table 6. Application of the test to series of nominal interest rates against the series of the inflation rate for sample sub-periods.

Series	<i>n</i>	<i>JREC</i>	<i>E(JREC)</i>	<i>JREC</i>	<i>E(JREC)</i>	<i>p-value</i>
Short term rate—Inflation rate	1P 231	380	265	1.43	1	0.0000
	2P 321	1348	513	2.62	1	0.0000
	3P 423	1601	893	1.79	1	0.0000
Medium term rate—Inflation rate	2P 321	1511	513	2.94	1	0.0000
	3P 423	1611	893	1.81	1	0.0000
Long term rate—Inflation rate	1P 411	1511	843	1.79	1	0.0000
	2P 321	2557	513	4.98	1	0.0000
	3P 423	1844	893	2.07	1	0.0000

would be of great interest, would require a much more detailed analysis. As we have previously mentioned, such an analysis lies beyond the scope of the present work.

6.2. The Relationship between Employment and GDP

This relationship is connected with the aggregate production function, which relates aggregate output to physical inputs or factors of production. Many theoretical and empirical works have been devoted to ascertaining the properties of the production function, and a broad agreement exists on some empirical relationships that this function must display. A widely accepted property of this function is that there is a nonlinear relationship between the quantities of productive factors and the amount of aggregate output obtained (*i.e.* that this function is nonlinear).

Next, we will apply our test procedure to check if this property is fulfilled in the US economy for labor input, and provide new empirical evidence. To do that, we will use the quarterly series of the real Gross Domestic Product (GDP), published by the US Bureau of Economic Analysis, and the series of the All Employees (total nonfarm), produced by the US Bureau of Labor Statistics⁹ (period: first quarter of 1960 to first quarter of 2015).

In the first row of **Table 7**, together with the results of our test, we also present the *p*-value of the Pearson correlation coefficient. This coefficient is a widely used tool to detect linear relationships between two variables. Its *p*-value (last column) is the probability of obtaining a correlation as large as the observed value by random chance, when the true correlation is zero.

We observe that both tests display the same result, both *p*-values are zero, which means the rejection of the null hypothesis for any significance level, and hence empirical evidence in favor of the existence of a relationship between GDP and employment.

In the second row, we present the results of the tests after filtering the linear component of the relationship. To do that, we fit a simple linear regression of the GDP on employment, and analyze the relationship between the residuals of this regression and employment. We observe that, in this case, the *p*-value of the correlation coefficient is equal to 1. This result is obvious, because if a regression model is computed correctly, the correlation coefficient between the residuals and the independent variable is zero.

However, the *p*-value of our test is still very small, so small as to reject the null hypothesis of independence for any usual significance level. Considering that one property of the regression models is that “the residuals have a nonlinear association with X if and only if the original observations of Y have a nonlinear association with X”, the conclusion will be that a nonlinear relationship exists between GDP and employment¹⁰.

This example is useful to show that our procedure can detect nonlinear dependence between variables, while other tests, such as the Pearson correlation coefficient cannot.

7. Conclusions

The main objective of this paper is to design a criterion that allows us to determine whether two series come

⁹To eliminate the trend, we will take both series in differences.

¹⁰In our case, the GDP is clearly the dependent variable (Y) while the Employment is the independent variable (X).

Table 7. Application of the test to the series GDP and Employment.

Series	<i>n</i>	<i>NJR</i>	<i>E(NJR)</i>	<i>JREC</i>	<i>p</i> -value	Corr. Coef. <i>p</i> -value
GDP-Employment	221	313	243	1.29	0.0000	0.0000
GDP-Employment (residuals of the LRM)	221	287	243	1.18	0.0010	1

from independent dynamical systems. We develop our procedure on the basis of the Joint Recurrence Plot, and more specifically on its main quantitative indicator, the percentage of recurrence points. This indicator measures the level of dependence or synchronization existing between the series. The values of this indicator have a clear interpretation when they are extreme (very high or very low), but they are not so obvious for intermediate values. To solve this problem, we derive the probability distribution of the variable “number of recurrence points in the JRP” under the null of independence. This (exact) probability distribution allows us to determine a cut-off point from which determines whether the series present statistical dependence or not.

The performance of our procedure is analyzed through a simulation study for a broad set of systems in which, a priori, the existence, or not, of independence is evident. We have also applied it to series affected by increasing levels of noise, and to a pair of coupled systems with different degrees of coupling. The obtained results are very satisfactory and confirm the efficacy of the proposed method.

Finally, we use this methodology to study the empirical relationship between the nominal interest rate and the inflation rate, and between the GDP and the employment. In both cases, the results are clear, in line with the consensus on these issues in the existing literature. Moreover, in this latter case we have shown that our test is capable of detecting both linear and nonlinear relationships.

Acknowledgements

This work was supported by the Spanish Ministries of Science and Innovation [ECO2009-07936] and Economy and Competitiveness [ECO2013-41353-P], and by the Government of Aragon [Consolidate Group S10, Consolidate Group S21].

References

- [1] Szekeley, G., Rizzo, M. and Bakirov, N. (2007) Measuring and Testing Independence by Correlation of Distances. *Annals of Statistics*, **35**, 2769-2794. <http://dx.doi.org/10.1214/009053607000000505>
- [2] Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M. and Sabeti, P.C. (2011) Detecting Novel Associations in Large Data Sets. *Science*, **334**, 1518-1524. <http://dx.doi.org/10.1126/science.1205438>
- [3] Heller, Y., Heller, R. and Gorfine, M. (2013) A Consistent Multivariate Test of Association Based on Ranks of Distances. *Biometrika*, **100**, 495-502. <http://dx.doi.org/10.1093/biomet/ass070>
- [4] Sklar, A. (1959) Fonctions de Répartition à *n* Dimensions et Leurs Marges. Vol. 8, Institut Statistique de l'Université de Paris, Paris, 229-231.
- [5] Siqueira, S., Takahashi, D.Y., Nakata, A. and Fujita, A. (2014) A Comparative Study of Statistical Methods Used to Identify Dependencies between Gene Expression Signals. *Briefings in Bioinformatics*, **15**, 906-918. <http://dx.doi.org/10.1093/bib/bbt051>
- [6] Eckmann, J.P., Kamphorst, S. and Ruelle, D. (1987) Recurrence Plots of Dynamical Systems. *Europhysics Letters*, **4**, 973-977. <http://dx.doi.org/10.1209/0295-5075/4/9/004>
- [7] Zbilut, J.P. and Webber Jr., C.L. (1992) Embeddings and Delays as Derived Quantification of Recurrence Plots. *Physics Letters A*, **171**, 199-203. [http://dx.doi.org/10.1016/0375-9601\(92\)90426-M](http://dx.doi.org/10.1016/0375-9601(92)90426-M)
- [8] Webber Jr., C.L. and Zbilut, J.P. (1994) Dynamical Assessment of Physiological Systems and States Using Recurrence Plot Strategies. *Journal of Applied Physiology*, **76**, 965-973.
- [9] Holyst, J.A., Zebrowska, M. and Urbanowicz, K. (2001) Observations of Deterministic Chaos in Financial Time Series by Recurrence Plots, Can One Control Chaotic Economy? *The European Physical Journal B*, **20**, 531-535. <http://dx.doi.org/10.1007/PL00011109>
- [10] Belaire-Franch, J. (2004) Testing for Non-Linearity in an Artificial Financial Market: A Recurrence Quantification Approach. *Journal of Economic Behavior & Organization*, **54**, 483-494. <http://dx.doi.org/10.1016/j.jebo.2003.05.001>

- [11] Kyrtsov, C. and Vorlow, C.E. (2005) Complex Dynamics in Macroeconomics: A Novel Approach. In: Diebolt, C. and Kyrtsov, C., Eds, *New Trends in Macroeconomics*, Springer-Verlag, Berlin, 223-238. http://dx.doi.org/10.1007/3-540-28556-3_11
- [12] Zbilut, J.P. (2005) Use of Recurrence Quantification Analysis in Economic Time Series. In: Salzano, M. and Kirman, A., Eds, *Economics: Complex Windows*, Springer-Verlag, Milan, 91-104. http://dx.doi.org/10.1007/88-470-0344-x_5
- [13] Barkoulas, J.T. (2008) Testing for Deterministic Monetary Chaos: Metric and Topological Diagnostics. *Chaos, Solitons and Fractals*, **38**, 1013-1024. <http://dx.doi.org/10.1016/j.chaos.2007.01.065>
- [14] Aparicio, T., Pozo, E. and Saura, D. (2009) Detecting Determinism Using Recurrence Quantification Analysis: Three Test Procedures. *Journal of Economic Behavior & Organization*, **65**, 768-787. <http://dx.doi.org/10.1016/j.jebo.2006.03.005>
- [15] Aparicio, T., Pozo, E. and Saura, D. (2011) Detecting Determinism Using Recurrence Quantification Analysis: A Solution to the Problem of Embedding. *Studies in Nonlinear Dynamics and Econometrics*, **15**, 1-10.
- [16] Aparicio, T., Pozo, E. and Saura, D. (2013) Do Exchange Rate Series Present General Dependence? Some Results Using Recurrence Quantification Analysis. *Journal of Economics and Behavioral Studies*, **5**, 678-686.
- [17] Kyrtsov, C. and Terraza, M. (2010) Seasonal Mackey-Glass-GARCH Process and Short-Term Dynamics. *Empirical Economics*, **38**, 325-345. <http://dx.doi.org/10.1007/s00181-009-0268-8>
- [18] Guhathakurta, K., Bhattacharya, B. and Roy Chowdhury, A. (2010) Using Recurrence Plot Analysis to Distinguish between Endogenous and Exogenous Stock Market Crashes. *Physica A: Statistical Mechanics and Its Applications*, **389**, 1874-1882. <http://dx.doi.org/10.1016/j.physa.2009.12.061>
- [19] Marwan, N. (2008) A Historical Review of Recurrence Plots. *European Physical Journal—Special Topics*, **164**, 3-12. <http://dx.doi.org/10.1140/epjst/e2008-00829-1>
- [20] Schinkel, S., Dimigen, O. and Marwan, N. (2008) Selection of Recurrence Threshold for Signal Detection. *The European Physical Journal—Special Topics*, **164**, 45-53. <http://dx.doi.org/10.1140/epjst/e2008-00833-5>
- [21] Zbilut, J.P., Giuliani, A. and Webber Jr., C.L. (1998) Detecting Deterministic Signals in Exceptionally Noisy Environments Using Cross-Recurrence Quantification. *Physics Letters A*, **246**, 122-128. [http://dx.doi.org/10.1016/S0375-9601\(98\)00457-5](http://dx.doi.org/10.1016/S0375-9601(98)00457-5)
- [22] Romano, M.C. (2004) Synchronization Analysis by Means of Recurrences in Phase Space. Ph.D. Dissertation, University of Potsdam, Potsdam.
- [23] Marwan, N., Romano, M.C., Thiel, M. and Kurths, J. (2007) Recurrence Plots for the Analysis of Complex Systems. *Physics Reports*, **438**, 237-329. <http://dx.doi.org/10.1016/j.physrep.2006.11.001>
- [24] Goswami, B., Ambika, G., Marwan, N. and Kurths, J. (2012) On Interrelations of Recurrences and Connectivity Trends between Stocks Indices. *Physica A: Statistical Mechanics and Its Applications*, **391**, 4364-4376. <http://dx.doi.org/10.1016/j.physa.2012.04.018>
- [25] Goswami, B., Marwan, N., Feulner, G., Kurths, J. (2013) How Do Global Temperature Drivers Influence Each Other? *The European Physical Journal—Special Topics*, **222**, 861-873. <http://dx.doi.org/10.1140/epjst/e2013-01889-8>
- [26] Grebogi, C., Ott, E., Pelikan, S. and Yorke, L. (1984) Strange Attractors That Are Non Chaotic. *Physica D: Nonlinear Phenomena*, **13**, 261-268. [http://dx.doi.org/10.1016/0167-2789\(84\)90282-3](http://dx.doi.org/10.1016/0167-2789(84)90282-3)
- [27] Packard, N., Crutchfield, J.P., Farmer, J.D. and Shaw, R.S. (1980) Geometry from a Time Series. *Physical Review Letters*, **45**, 712-716. <http://dx.doi.org/10.1103/PhysRevLett.45.712>
- [28] Takens, F. (1981) Detecting Strange Attractors in Turbulence. In: Rand, D. and Young, L., Eds, *Dynamical Systems and Turbulence*, Springer-Verlag, Berlin, 366-381. <http://dx.doi.org/10.1007/bfb0091924>

Appendix: A Brief Description of the Dynamic Systems Employed

This Appendix provides details of the equations that generate the dynamic systems used in our study.

Henon System:

$$x_{i+1} = 1 - 1.4 \cdot x_i^2 + y_i$$

$$y_{i+1} = 0.3 \cdot x_i$$

Ikeda System:

$$t = 0.4 - \frac{6}{1 + x_i^2 + y_i^2}$$

$$x_{i+1} = 1 + 0.9 \cdot [x_i \cdot \cos(t) - y_i \cdot \sin(t)]$$

$$y_{i+1} = 0.9 \cdot [x_i \cdot \sin(t) + y_i \cdot \cos(t)]$$

GOPY System:

$$x_{i+1} = 2.8 \cdot \tanh(x_i) \cdot \cos(y_i)$$

$$y_{i+1} = y_i + 2 \cdot \pi \cdot \frac{\sqrt{5} - 1}{2}$$

Henon - Ikeda Coupled System:

$$x_{1,i+1} = 1 - 1.4 \cdot x_{1,i}^2 + y_i$$

$$x_{2,i+1} = 0.3 \cdot x_{1,i}$$

$$t = 0.4 - \frac{6}{1 + x_{3,i}^2 + x_{4,i}^2}$$

$$x_{3,i+1} = 1 + 0.9 \cdot [x_{3,i} \cdot \cos(t) - x_{4,i} \cdot \sin(t)] + \varepsilon \cdot x_{1,i}$$

$$x_{4,i+1} = 0.9 \cdot [x_{3,i} \cdot \sin(t) + x_{4,i} \cdot \cos(t)]$$

As can be seen, in the unidirectionally-coupled Henon - Ikeda System, the Henon and Ikeda systems behave respectively as *drive* and *response*. Specifically, the variable x_1 of the Henon system is coupled into the second equation of the Ikeda system, with ε being the “coupling strength”. Obviously, for $\varepsilon = 0$, drive and response would be independent.