Detecting Causal Relations Among Symbolic Time Series*

Fernando Delbianco[2], Federico Contiggiani[1], Andrés Fioriti[2], and Fernando Tohmé[2]

[1] *Universidad Nacional de Río Negro*
[2] *Universidad Nacional del Sur—INMABB, CONICET*

Abstract

Symbolic time series analyses (STSA) are used in economics and other social sciences as a way of reducing the impact of noise on data and to exhibit more clearly the evolution of time series. We show that causality tests applied to symbolic series may fail to detect actual relations or generate *statistical artifacts*. Well-known causality detection methods, like transfer entropy, Granger's test, or PCMCI, may miss some existing causal relationships or, more frequently, yield non-existent ones. The performance of these methods may differ, depending on the specific choices of lag structures and alphabet sizes, as well as on the characteristics of the underlying dynamic process.

Keywords: Causality, Symbolic time series analysis, SAX, Markov switching model, OPTN, Transfer entropy, Granger's test, PCMCI.

Our work reveals a vulnerability in analyzing complex systems. We show that causality detection techniques applied to symbolic time series (STS) can create spurious causal relationships or, less frequently, fail to detect actual ones. This research shows that transforming numerical data into a symbolic representation may affect the performance of causality detection methods, such as transfer entropy, Granger's test, or PCMCI, leading to the incorrect identification of causal links. The originality of our contribution lies in showing how the symbolization processes interact with causality detection procedures. Our work calls for a reassessment of current methodologies, advancing the field by developing more robust techniques for inferring causality in complex economic and social systems using symbolic data.

*Contiggiani: Avenida Leloir and Don Bosco, Viedma, Argentina. Delbianco, Fioriti and Tohmé: Department of Economics and INMABB-CONICET, San Andrés 800, 8000 Bahía Blanca, Argentina. Contiggiani: fcontiggiani@unrn.edu.ar, Delbianco: fernando.delbianco@uns.edu.ar (corresponding author), Fioriti: andres.fioriti@uns.edu.ar, and Tohmé: ftohme@criba.edu.ar.

1 Introduction

Time series analysis is a relevant tool for the study of dynamic systems. A central insight advanced by Hermann Wold and John von Neumann is that a time series can be decomposed as a *trend* plus a purely random *noise*. It is based on the assumption that the trend obeys to some deterministic process that, if it remains invariant over time, yields a *stationary* time series under an also invariant distribution of noises (Pourahmadi, 2001). Many of the methods of time series analysis involve filtering out the noise of the series and applying procedures to obtain a stationary variant that facilitates forecasting future values or to detect relations among different variables, in particular *causal* ones (Hamilton, 1994).

In certain applications of economic interest, the noise component is considered to be particularly confounding, hiding the regularities that guide the evolution of the variables of interest. Many techniques have been applied to address this question, mostly of a statistical nature. Alternatively, in certain analyses, the noise reduction procedure is qualitative. More precisely, the whole range of a series is replaced by a reduced "vocabulary", in such a way that its evolution consists of a trajectory through a small class of states (Brida and Punzo, 2003).

In the latter case, the ensuing *symbolic time series* allows a more straightforward application of the usual time series methods to infer relevant information. The implicit assumption underlying the use of this qualitative approach is that it works by eliminating obfuscating noise and making explicit the essential behavior of the system.

The translation of noisy and complex behavior into a simple narrative of stylized states and their transitions is indeed attractive. But it can also be misleading, given its potential for inducing false interpretations of the actual behavior of the systems under analysis.

The detection of causal relationships in time series is a cornerstone technique in understanding dynamic systems, particularly within economics and the social sciences. Traditionally, methodologies such as Granger causality (Granger, 1969) have been central for this purpose, especially when raw numerical data can be assumed to follow stationary or near-stationary processes. However, the transformation of numerical series into symbolic representations—via symbolic time series analysis (STSA) or related symbolic dynamics methods—has emerged as a popular approach to mitigate noise and highlight system regularities. As said, while this symbolization supposedly simplifies complexity, it can also be compounded with inherent limitations of methods of causality detection to yield wrong results. Spurious causal links may become accepted or genuine ones be discarded, yielding artifacts not present in the underlying data.

Several studies have proposed robust means to refine and reinterpret symbolic time series methods. Complexity measures based on ordinal partition transition networks (OPTNs), which complement classical symbolic and entropy-based methods, address the problem of making symbolic time series analyses more robust (Ruan et al., 2019). Their approach—rooted in complex networks theory—distills the evolution of ordinal patterns between multiple time series, allowing directionality and delay of coupling to be inferred directly from data. OPTNs offer information-theoretic indices (such as conditional entropy and Kullback-Leibler divergence) to capture the structure of couplings, supporting causal inference even in multivariate or nonlinearly coupled systems. The practical strength of OPTN-based complexity measures is shown in their success at diagnosing causality and interaction delays in both synthetic and real-world (e.g., climatological) data, addressing key limitations of pairwise, static, or linearly-assumed techniques relied upon in earlier STSA. An extension, based on the development of noise-resistant ordinal partitioning, leverages entropy-based criteria to select effective sections in the relationship between consecutive data points in time series (Shahriari et al., 2023).

At the same time, the broader landscape of weak signal extraction—while outside the primary STSA literature—offers lessons in the application of symbolic and noise-resistant analytics. A recent contri-

bution (Song et al., 2023) presents a quantum metrology-inspired method for extracting weak signals in non-stationary channels, incorporating weak measurement for real-time, high-sensitivity detection. Although their primary applications lie in quantum sensing and signal processing, their results highlight the benefits of advanced filtering (finite impulse response with LMS updating) and adaptive spectrum shifting. Such techniques parallel the need in symbolic time series analysis for robust preprocessing and careful noise separation, suggesting a convergence in methodological priorities across disciplines engaged in weak signal detection.

In this paper, we explore how different methods of symbolization and causality detection in symbolic time series may affect the results. More precisely, we seek to understand under which conditions causal detection methods interacting with symbolic methods can fail to detect relevant features in the original data. Users of these approaches may end up discovering *statistical artifacts*, instead of real phenomena. The results of analyzing symbolic time series may be produced by how the symbolic translations affect the causality inference methods rather than by the phenomenon being studied (Vogt and Johnson, 2015). This is particularly problematic in the case of applications of a social or economic nature, since this qualitative method is usually applied in those fields to detect patterns that may be useful not only to understand actual processes but to recommend policies.

The choice of the number of parts in which the range of values is partitioned is equivalent to selecting a particular *dictionary* of symbols, translating an original real-valued series into a symbolic time series. The question is then, what is the impact of the choice of dictionary on the results of the time series analysis? We show that the answer in the case of the detection of causal relations is that the choice of symbolic time series may contribute to the detection of non-existing ones or the omission of some existing ones. But some other methods may help to ameliorate this negative effect.

In the rest of this paper, we will discuss this issue in detail. Section 2 presents a discussion of the different ways in which symbolic time series are obtained from series of actual numerical data. Section 3 discusses three tests of causal detection in time series. In Section 4, we run three simulations and apply symbolic transformations of the data to show when causality detection tests may fail on the resulting series. We further illustrate this in a real-world case in which the data generation process is unknown. Section 5 concludes.

2 Symbolic Time Series

Symbolic time series analyses (STSA) have been used in economics as an alternative to the usual time series analyses in econometrics. The key idea is that the data generated by a dynamical system and recorded according to its associated measurement scale is translated into a symbolic representation. This translation does not assume that the underlying system verifies any specific structural or stationarity condition. The promise of this framework is the possibility of reconstructing the data, capturing the essential properties of the generating system, while giving the researcher more freedom to select an appropriate symbolic coding.

This approach has been used in investigations on business cycles, where the purpose is finding well-adjusted models capturing the breakpoints of series and replicating the dynamics of the underlying systems (Brida and Punzo, 2003; Brida, 2006; Brida et al., 2009; Chao, 2011). A related research field in economics to which STSA has been applied is growth theory, particularly in cases where the objective is to study disparate evolution paths of economies or economic sectors (Brida and Punzo, 2003).

The next subsections cover the methods of symbolic translation of numeric time series that will be used in our empirical analyses. They may benefit from the application of pre-filtering and denoising procedures before the application of the proper symbolization method. In the case of signal processing, this is a

fundamental component of the analysis (see Song et al. (2023) in the case of quantum information), and more generally, when a time series shows high variance, outliers, or abrupt level shifts not attributable to the system dynamics of interest. In our analyses, such procedures are either not necessary (in the case of simulated series) or already incorporated in the series (see Delbianco et al. (2021)).

2.1 Symbolic aggregate approximation (SAX)

Lin et al. (2003) applies a variant of STSA known as *Symbolic aggregate approXimation* (SAX), in which the dimensionality of a time series is reduced by using two kinds of transformations. The first one is *piecewise aggregate approximation* (PPA), which replaces the original series by a sequence of segments corresponding to the average of the values in a time bin, specified arbitrarily. The second procedure involves translating the PPA series into a discrete string of symbols drawn from a given dictionary.

SAX can be described summarily as follows:

- *PPA step*: a time series $C = \{c_j\}_{j=1}^{n}$ of length n is represented by a vector $\bar{C} = (\bar{c}_1, \ldots, \bar{c}_w)$ in a w-dimensional space.

  The ith element of $\bar{C}$ is calculated as:

$$\bar{c}_i = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} c_j \tag{1}$$

- *Symbolization step*: an alphabet size is defined as an arbitrary integer $d > 2$. The PPA time series is normalized and a sorted list of *breakpoints* $B = \beta_1, \ldots, \beta_{|d|-1}$ is defined such that the area under a $N(0,1)$ Gaussian curve from β_i to β_{i+1} is $\frac{1}{|d|}$. Once defined a list of symbols $\{\alpha_i\}_{i=1}^{d}$, the PAA approximation $\bar{C}$ is translated into a *word* $\hat{C} = \hat{c}_1, \ldots, \hat{c}_w$, where $\hat{c}_i = \alpha_j$, iff $\beta_{j-1} \leq \bar{c}_i < \beta_j$.[1]

This process transforms a continuous series into a symbolic one. The sections that specify the new series are defined according to the chosen dictionary (the number of sections to be defined in STSA) and the original distribution of the data.

2.2 Markov switching models

A *Markov switching model* (MSwM) allows us to characterize how a non-stationary series transitions between different states, by centering on the probability distribution of the switches between those states (Hamilton, 1989, 1990). In an MSwM, a *state* is equivalent to a symbol in STSA. In this sense, it yields another way of reducing the dimensionality of a time series.

Consider a system transitioning between a finite number of states $1, \ldots, N$, such that for any period $t \in \mathbb{N}$ the distribution of possible instances of the state variable s_t satisfies the following condition:

$$P\{s_t = j | s_{t-1} = i, s_{t-2} = k, \ldots\} = P\{s_t = j | s_{t-1} = i\} = p_{ij} \tag{2}$$

where each p_{ij} corresponds to the probability of a transition from state j to state i and then, $p_{i1} + p_{i2} + \cdots + p_{iN} = 1$.

Now consider a time series $\{y_t\}_{t\geq 0}$. This means, in particular, that if the values are drawn from a compact set Y, the distributions over Y at t, F_t, and at any period $t + k$, F_{t+k}, do not necessarily

[1] An alternative symbolization procedure is based on segmenting the range of data values in quantiles, determined by their distribution.

verify that $F_t = F_{t+k}$. The behavior of the series can be described as follows (Hamilton (1989), Hamilton (1994)):

$$y_t - \mu_{s_t^*} = \phi\left(y_{t-1} - \mu_{s_{t-1}^*}\right) + \varepsilon_s \tag{3}$$

where $\mu_{s_t^*} = E[Y_t \mid s_t^*] \in Y$ corresponds to the state $s_t^* \in \{1, \ldots, N\}$. If $s_t^* = j$ and $s_{t-1}^* = i$, at $t-1$, μ_i is followed in t by μ_j, with $\mu_i \neq \mu_j$. The transition from $\mu_{s_{t-1}^*}$ to $\mu_{s_t^*}$, corresponding to transition from state j to state i has probability p_{ij}. Thus ϕ is a function that embodies the combined action of P and, for each state i and period t, the conditional distribution $F_t(y|i)$.

Then, any time series can be reduced to a sequence of states, each of which drives the evolution of the underlying system according to its corresponding mean value.

In our study we focused on regime changes in the mean value and applied the default convergence criteria of the R language MSwM package (Sanchez-Espigares and Lopez-Moreno, 2021), which utilizes the *expectation maximization* (EM) algorithm. We report only those cases in which convergence has been achieved, while those in which it failed were omitted. The number of regimes has not been selected arbitrarily, focusing on models with two, three, and four regimes. The default highest number of iterations was set to 100.

2.3 Ordinal partition transition networks (OPTN)

An alternative framework (Ruan et al., 2019) generates symbolic time series via *ordinal partition transition networks* (OPTN). Here, nodes represent all $m!$ possible ordinal patterns (where m is the length of windows extracted from a time series), and directed, weighted edges record the empirical frequencies of transitions between consecutive ordinal patterns along the time series:

$$A_{ij} = \frac{\text{number of transitions from ordinal pattern } i \text{ to } j}{\text{total transitions from } i}.$$

The resulting symbolic time series is the sequence of ordinal patterns assigned to each window, and statistical descriptors (degree distributions, entropy, Kullback-Leibler divergence) of the transition network supply a symbolic complexity profile of the dynamics. Importantly, if studying two (possibly coupled) time series X and Y, constructing a *bipartite* OPTN—with transitions $i \to j$ only if pattern i in X is followed, with delay Δt, by pattern j in Y—allows one to analyze directed interactions and infer coupling direction and delay by maximizing measures such as conditional entropy or divergence.

The procedure is as follows:

1. Select ordinal window size m.
2. For each $i = 1, \ldots, N - (m-1)$, extract window $\mathbf{z}_i$ and assign an ordinal symbol sequence corresponding to the rank ordering of elements.
3. Assemble symbolic time series as the sequence of assigned ordinal symbols.
4. (For OPTN) Construct a transition matrix evaluating the transitions between ordinal symbols.

The goal is to obtain a fine-grained symbolic representation rooted in an ordinal structure, which can reveal system invariants even from short or contaminated data.

2.4 Approaches to the symbolization of time series

There exist different approaches to the translation of numerical time series into symbolic representations. We can distinguish some key methodologies. Some are discussed in this paper, while others are presented just for the sake of completeness:

- *Piecewise aggregate approximation (PAA) + SAX (symbolic aggregate approximation) (Lin et al., 2003, 2007).* PAA reduces noise by taking the means of the numerical data over fixed-length windows. This decreases the number of values to consider while preserving the trend of the time series. SAX translates those average values into discrete symbols according to designated mapping rules. This procedure is computationally efficient and preserves trends, but conditional on assuming normal distributions of the numerical values. Furthermore, the result is sensitive to the size and placement of the windows.

- *Quantile-based binning (Daw et al., 2003).* This procedure divides the value range into equiprobable bins using empirical quantiles (e.g., terciles, quartiles), providing results robust to outliers, respecting the distribution of numerical data. This comes at the price of ignoring the temporal order of the observations, creating a noisy representation in the case of non-stationary series.

- *Model-based methods.* These approaches assume an underlying model of the data-generating process to detect patterns and assign symbols. One of these methods uses the algebraic-topological concept of persistent homology to identify significant peaks and valleys, assigning symbolic representations to them (Seversky et al., 2016; Bois, 2024). The main strength of these methods resides in their ability to capture features present at different scales. On the other hand, their main limitations are their sensitivity to the choice of parameters and their high computational cost. Another model-based method is the Markov-switching approach used in this paper, which assumes a probabilistic state-space model with discrete hidden states that switch according to a Markov chain. The observed time series is discretized into symbols by assigning each time point to the most likely hidden state (via Viterbi decoding). The symbol alphabet can be either learned from data or fixed a priori (Hamilton, 1989, 1990).

- *Shape-based methods (Malinowski et al., 2013)).* These procedures discretize local trends (up, down, steady) using first and second derivatives or differences, providing results robust to scaling. Their downside is their sensitivity to noise if no previous smoothing procedures have been applied.

- *Clustering-based methods (Lonardi and Patel, 2002).* These procedures partition in clusters the observations, assigning symbols to each cluster. While they uncover natural patterns ("motifs"), they require a predefined number of clusters and are computationally expensive.

PAA + SAX and quantile-based binning are bin-based methods, in which the key element is the choice of the bins of values to be assigned symbols. Markov switching is, as said, model-based, using the underlying states of a Markov process to assign symbolic values. Shape-based methods are ordinal, ranking local trends according to derivatives (or differences, in the case of discrete time series). Clustering-based methods are network-pattern based, looking for clusters across the series to assign symbols representing that distributed information.

OPTN, another method used in this study, builds a complex network where nodes are identified with ordinal patterns (permutation ranks of local subsequences) and edges with observed transitions between them. The symbolic sequence is derived from the ordinal ranking of values within a sliding window. This network captures higher-order dynamics and overlaps with clustering-based motif discovery, but with an emphasis on the graph structure over cluster centroids Small (2013).

Each of these methods has a preferred condition for its applicability. So, for instance, in the case of high non-linearity, either PAA + SAX or the shape-based approach can capture more easily the varying trends. Quantile-based binning is preferable in the case of stationary series, while shape-based or persistence-based methods seem more appropriate to detect non-stationarity, particularly because of their robustness to scaling.

With respect to the response to the presence of noise, quantile binning and clustering perform better with little noise. SAX with larger PAA windows, shape-based methods after preliminary smoothing, and persistence-based methods are robust to higher degrees of noise in the data.

In summary, bin-based methods work better with stationary data, while ordinal ones provide good trend detection capacities in noisy or non-stationary regimes. Model-based approaches are more appropriate for capturing multi-scale patterns, and network-pattern methods can be recommended for detecting patterns in repetitive systems.

3 Causality analysis of time series

In econometrics, the concept of *causality* is customarily identified with the existence of *Granger*-causality relations among time series (Granger, 1969). The latter concept captures the idea that a variable x "causes" a variable y if its current and past values provide better explanations of the future value of y than only the past values of y. This notion can be rather easily tested if the relation between the variables is assumed to be linear.

Alternatively, the existence of causal relations among time series can be understood in terms of the transfer of entropy from one series to another. A test detects the amount of information that a series corresponding to a variable x transfers to a series corresponding to y.

Another approach analyzes the potential causal graph among the variables of interest. Testing for conditional independence allows for removing irrelevant nodes until only the causal relations remain.

The following three subsections describe more precisely the tests that establish whether a relation among series satisfies the conditions of any of these three notions of causality.

3.1 Granger causality

Let $\{X_t\}_{t\geq 0}$ and $\{Y_t\}_{t\geq 0}$ be two stationary time series and the following *vector autoregression* (VAR) model, relating them:

$$Y_t = \delta + \sum_{j=1}^{p} \theta_{11,j} Y_{t-j} + \sum_{j=1}^{p} \theta_{12,j} X_{t-j} + u_{Yt} \tag{4}$$

$$X_t = \eta + \sum_{j=1}^{p} \theta_{21,j} Y_{t-j} + \sum_{j=1}^{p} \theta_{22,j} X_{t-j} + u_{Xt} \tag{5}$$

X is said to cause Y if X_t "Granger-causes" Y_t but not the other way around:

$$E(Y_t \mid Y_{t-1}, \ldots, Y_{t-p}, X_{t-1}, \ldots, X_{t-p}) \neq E(Y_t \mid Y_{t-1}, \ldots, Y_{t-p}) \tag{6}$$

but

$$E(X_t \mid X_{t-1}, \ldots, X_{t-p}, Y_{t-1}, \ldots, Y_{t-p}) = E(X_t \mid X_{t-1}, \ldots, X_{t-p}) \tag{7}$$

In terms of the VAR model, this is equivalent to stating that the estimated coefficients are such that $\boldsymbol{\theta_{12,j}} \neq 0$ while $\boldsymbol{\theta_{21,j}} = 0$.

3.2 Transfer entropy (TE)

If no underlying linear relation is assumed between X and Y, transfer entropy helps to detect causal influences among them (Schreiber, 2000; Hlaváčková-Schindler et al., 2007).

Although there are several entropy definitions, we will focus on Rényi's concept of entropy since its formula generalizes other measures:

$$H_q = \frac{1}{1-q}\log\left(\sum_{i=1}^{n} p_i^q\right) \tag{8}$$

As $q \to 1$, Rényi entropy converges to Shannon's entropy:

$$H(X_i) = -\sum_{i=1}^{n} p_i \log(p_i) \tag{9}$$

To test the gain in information of a time series Y_i knowing the realization of a time series X_i, a function $TE(X \to Y)$ is defined as follows:

$$TE(X \to Y) = H(Y_t \mid Y_{t-1:t-p}) - H(Y_t \mid Y_{t-1:t-p}, X_{t-1:t-p})$$

The corresponding causality test is carried out under the null hypothesis that $TE(X \to Y) = 0$. If it is rejected, X is said to cause Y.

In our exercises we employ the `RTransferEntropy` package (Simon et al., 2019), with Rényi entropy with two values of the order parameter (q): 0.5 and 0.9. Each value responds to a slightly different goal: when $q = 0.5$ the measure places a greater emphasis on rare or low-probability events, whereas $q = 0.9$ yields a more uniformly distributed weighting across all events. We use a temporal window corresponding to a single lag, and the probability densities are estimated via kernel density estimation.

3.3 PCMCI

The PCMCI (Runge et al., 2019) method is based on conditional independence testing. It employs partial correlation tests, estimating these correlations based on the residuals obtained from a linear regression between the variables of interest.

It is based on the PC algorithm (Spirtes, 1991), which starts with a fully connected, undirected graph and progressively removes edges by testing for conditional independencies. *PCMCI* (Runge et al., 2019) operates in two stages: (i) it first applies a time series adaptation of PC to identify potential parent nodes for each variable at time t, denoted by X_t^j, and (ii) it then uses the *momentary conditional independence* (MCI) test to examine whether $X_{t-\tau}^i \to X_t^j$, where $X_{t-\tau}^i$ is a time-lagged parent of X_t^j. The first stage thus constructs the *Markov* set for each node by removing irrelevant variables, while the second stage eliminates false positives from the identified relationships.

We used the default PCMCI test drawn from the TRIGAMITE package [2], specifically the ParCorr (partial correlation) variant, to assess independence. For the lag structure we considered lags ranging from 0 to 3. The null lag is reported as the contemporaneous effect, while among the positive lags we report only the smallest one that yields a statistically significant result, if any. The significance level (α) for the PC algorithm is left at its default value of 0.05. Multiple-testing correction uses the Benjamini–Hochberg procedure to control the false discovery rate.

[2] `https://github.com/jakobrunge/tigramite`

4 Exercises with data

Having presented the main concept of STSA and three important ways of detecting causality relations in time series, we are now ready to see the results of the interaction of symbolization with the application of the latter. We will see that in controlled settings (simulated causal relationships), the replacement of numerical values by symbols may contribute to the detection of relations that are actually absent in the original formulation. We run three simulation exercises, with a linear, a non-linear, and a random walk representation, respectively, obtaining different results.

Then, we focus on real-world data in a case in which it is relevant to detect whether a causal relation exists and, if so, to determine its direction. We will see that the choice of symbolic transformation, the causality detection method, and even the use of differentiation of data may affect the results of the analysis.

4.1 Simulation of a linear causal relation

The first exercise we run considers three time series systems between which we define a linear relation. More precisely, let the three time series, $X_{1,t}$, $X_{2,t}$ and $Y_{1,t}$, be defined by the following data generation processes:

$$X_{1,t} = \delta_1 + \sum_{i=1}^{6} \theta_{1,i} X_{1,t-i} + u_{1,t} \tag{10}$$

$$X_{2,t} = \delta_2 + \sum_{i=1}^{6} \theta_{2,i} X_{2,t-i} + u_{2,t} \tag{11}$$

and

$$Y_{1,t} = \gamma + \phi_1 X_{1,t-1} + v_t \tag{12}$$

where:

- $(\theta_{1,1}, \theta_{1,2}, \ldots, \theta_{1,6}) = (0.3, 0, 0, 0, 0, 0.6)$.
- $(\theta_{2,1}, \theta_{2,2}, \ldots, \theta_{2,6}) = (0.3, 0, 0.3, 0, 0, 0.3)$.
- $\phi_1 = 0.5$, $\delta_1 = 0$, $\delta_2 = 1$, $\gamma = 0$
- $u_{1,t}, u_{2,t}$ and v_t follow a Normal distribution with mean 0 and variance 1.

In many cases, for instance, when series are non-stationary (not the case here), it is necessary to differentiate them. Thus, the series of interest for analysis is that of the *differences* of the original ones. To represent this, in our case, we also consider the following specifications:

$$\Delta X_{1,t} = \sum_{i=1}^{6} \rho_{1,i} X_{1,t-1-i} + w_{1,t} \tag{13}$$

$$\Delta X_{2,t} = \sum_{i=1}^{6} \rho_i X_{2,t-1-i} + w_{2,t} \tag{14}$$

$$\Delta Y_{1,t} = \psi_i X_{1,t-1} + z_t \tag{15}$$

We can now run the causality detection tests on them. We test for the existence of Granger's causality, TE, evaluating Rényi's entropy at q=0.5 and q=0.9, and PCMCI.

We use SAX (in both the Gaussian and quantile versions, denoted with g and q, respectively), the Markov switching model, and OPTN. For SAX and Markov, we use dictionaries with 2, 3, and 4 symbols, while for OPTN we specify $M = 3$ and $\tau = 5$ for every simulation. We will show the results of the PCMCI tests in two columns, one with the contemporaneous effect (lag 0), and another one for any positive significant lag ($lag > 0$). More specifically, this second column summarizes the detection of relations for lags 1, 2, etc.

The tables in the appendix presenting the results of different simulations show the p-values of the estimations of the relations between the corresponding variables.[3] The first row shows the results when the original series is used, while the other rows correspond to different symbolization methods.

The columns, in turn, represent the causality tests under different conditions. Each table corresponds to the tests of a particular causal direction. In parentheses, we report the results for the opposite direction. The results of the Table 1 obtained with the original series under Granger's test (with two lags), TE (with $Q = 0.9$), and PCMCI (with positive lag) are highly significant, capturing the specified causal relation $X_1 \to Y_1$, as defined by equations (10)-(12). This is also the case, under the same specifications for all the methods of symbolization, indicating that they faithfully capture the causal relation.

Table 2 indicates that the non-causal relation $Y_1 \to X_2$ is wrongly detected by TE (for both $Q = 0.5$ and $Q = 0.9$) and PCMCI (without lags). Interestingly, Gaussian SAX with 2 and 3 symbols, quantile SAX with 3 and 4 symbols, as well as OPTN do not yield a false positive.

In tables 3 and 4, the results for the opposite directions present interesting contrasts. The acceptance of both $Y_1 \to X_1$ and $X_2 \to Y_1$ would yield false positives, but in the latter case, only TE (with $Q = 0.5$) and PCMCI (with positive lags) can reject it, even with the original series. In the former case, only Granger fails to do so with the original data. Under symbolization, this is somewhat turned around, with more cases of rejection of the false causal relation in the case of $X_2 \to Y_1$ than in that of $Y_1 \to X_1$.

Taking differences yields systematically more accurate results for all four possible causal relations between the Xs and Y_1, both with the original data (except one case in which PCMCI without lags fails) and with all the symbolization methods, other than a few exceptions.

4.2 Simulation of a non-linear causal relation

As a second exercise, consider two time series, X_t and Y'_t, related by the following data generation process, a polynomial of order 3:

$$Y'_{1,t} = \gamma + \phi_{2,i} X_{1,t-1} + \phi_{3,i} X^2_{1,t-1} + \phi_{4,i} X^3_{1,t-1} + v_t \tag{16}$$

where:

- $X_{1,t}$ is the same variable of our previous exercise, and the parameters of its specification are the same as in the linear case.
- $\phi_2 = 0.75$, $\phi_3 = 0.5$, $\phi_4 = 0.25$
- $v_t \sim \mathcal{N}(0, 1)$.

Table 5 presents the result of testing for the actual causal relation $X_1 \to Y'_1$. We can see that it fails to be detected only by PCMCI without lags, for all the symbolization methods except for Gaussian SAX with 2 symbols. But this can be expected, since the series is based on lags of X_1, and PCMCI without lags should fail in detecting the relation. Differentiation improves, slightly, the results obtained with the original data. Interestingly, as shown in Table 6, quantile SAX with 2 symbols accepts $X_1 \to Y'_1$ with

[3] In all the tables, ns means "not significant", $*$ significant at 5%, $**$ at 1%, while $***$ corresponds to $p \leq 0.001$.

high significance in the case of PCMCI with no lags, but fails to do so for the same causality criterion when positive lags are included in the analysis.

Table 5 presents in parenthesis the results for the non-causal relation from Y_1' to X_1. All the causality tests reject it when no symbolization is used. Instead, Gaussian SAX with two symbols fails to reject it, except in the case of Transfer entropy (with $Q = 0.5$). Another interesting observation is that Granger causality with two lags fails to reject the wrong relation, for every choice of symbols. Differentiation, as shown in Table 6, unlike in the case of the linear model, *worsens*, albeit slightly, the results.

4.3 Simulation of a random walk

As a final exercise, we generate a random walk series to analyze causal detection with symbolic non-stationary series.

With the same specification of $X_{1,t}$, ϕ_i $(i = 2, 3, 4)$, σ, and v_t as in the previous simulations, in this scenario, a variable $Y''_{i,t}$ follows a *random walk*:

$$Y''_{1,t} = \gamma + Y''_{1,t-1} + \mathcal{N}(\phi_{2,i} X_{1,t-1} + \phi_{3,i} X^2_{1,t-1} + \phi_{4,i} X^3_{1,t-1}, \sigma) + v_t \tag{17}$$

We include $X_{2,t}$, defined as in the case of the linear specification, to check whether our procedures correctly detect its independence from Y''_1.

Table 7 shows that the causal detection methods that yield accurate answers for symbolic versions of the series are fundamentally Granger and PCMCI with positive lags, except in the case of Markov with 2 symbols.

In the case of the non-causal relation from Y''_1 to X_1, both Granger and PCMCI with positive lags *fail* to reject it with the original values of the series, but other than that, with symbolic series, all the causality detection (except Granger with OPTN) methods correctly discard this relation.

Table 8 indicates the results for the independent variables X_2 and Y''_1. We can see that, with only a few exceptions, most notably stemming from the TE test and Granger and PCMCI with positive lags for OPTN, statements of causality between these variables are rejected.

Differentiating the three series, we find that (Table 9) Granger with 2 lags and PCMCI with positive lags, detect the causal relation $X_1 \to Y''_1$ with high significance, both in the numerical data as well as in all the symbolic variants. These causality detection methods also perform well, this time rejecting $Y''_1 \to X_1$, except in the case of the Markov symbolization with 3 terms in the dictionary.

Finally, as indicated in Table 10, all the causality detection methods for almost all the symbolizations can detect the absence of a causal relation between the independent variables X_2 and Y''_1.

4.4 Continuous dynamics

We consider now a system in which x_1 causes y_1 and vice versa, while a variable x_2 evolves autonomously, described by differential equations. One of those equations, a first-order differential equation, characterizes a non-conservative, oscillating system with non-linear damping (a van der Pol-like oscillator). The other two equations are linear differential equations. More precisely:

$$\frac{dy_1}{dt} = x_1$$
$$\frac{dx_1}{dt} = \mu\left(1 - y_1^2\right) x_2 - x_1$$
$$\frac{dx_2}{dt} = \alpha\, x_2$$

The parameters are: $\mu = 5$ and $\alpha = 0.05$, and the initial conditions: $x_1 = 2, y_1 = 0, x_2 = 1$. We also add noise that follows a normal distribution $\mathcal{N}(0, \sigma)$ with $\sigma = 0.1$.

We generated 10000 data points, sampling 1000 at regular intervals, to ensure the comparability with the previous simulations.

In general, all the methods applied to the series of x_1, x_2, and y_1 make errors but also capture some of the phenomena well. Tables 11 to 14 present the results.

We can see that transfer entropy is the method with the best overall performance, by primarily capturing the two-way relations between x_1 and y_1, without yielding many false positives between y_1 and x_2. Granger does not report many false positives between y_1 and x_2, but yields some on the relationship between x_1 and y_1, detecting usually only the one from y_1 to x_1 but not the opposite direction. PCMCI, on the other hand, fails to capture any relationship between y_1 and x_1 at lags 0 and 1, usually pointing to x_1 as the cause of y_1. Regarding false positives, it spuriously detects that y_1 causes x_2 in some instances. As for the symbolic series, no single one is clearly superior. They tend to be more dependent on the interaction with the method of detection of causal direction.

4.5 False positives and negatives

The tables detailing the results of the simulation exercises show when errors of types I and II are present, indicating the presence of both false positive and false negative causal relations. A general pattern emerges, namely that both kinds of errors happen less frequently when the series are differentiated, previous to obtaining a symbolic variant and/or testing for the presence of causal relations.

Errors of type I happen when the null hypothesis is mistakenly rejected as a result of the test procedure. In our analysis, it means accepting a non-existent causal relationship. In the case of the linear model, where the only causal relationship is $X_1 \to Y_1$, we can see that it is more common when the causal detection method is transfer entropy or Granger, when the series is not differentiated. OPTN also contributes in the case of PCMCI.

In the non-linear case, type I errors arise more frequently with Granger, even in the case of differentiated series. This is not surprising, since it was conceived for linear relations. Interestingly, with differentiated data, the performance of OPTN worsens, at least when the causality test is based on PCMCI.

In the random walk simulations, while no clear pattern emerges in the case in which the original values of the series are kept, differentiation does not eliminate errors: in the case in which Y_1'' is tested as a cause of X_1, all the causality tests fail at some point, except when Markov switching with 3 regimes provides the symbolization.

Type II errors happen when a true causal relationship fails to be detected. For our linear model, this happens with PCMCI with no lags, particularly in the differentiated case. It also happens with other causality tests when SAX and Markov switching are the symbolization methods applied.

For the non-linear model case, there is only one instance of type II error. It involves the use of PCMCI without lags as the causality test. Differentiation only contributes marginally to improve the

PCMCI with 0 lags performance and quantile SAX with two terms.

Finally, in the case of the random walk model, TE in its two versions and PCMCI with 0 lags do have a dismal performance, while OPTN as a symbolization method outperforms SAX and Markov switching. In differences, these behaviors improve, although Markov switching with three regimes has the worst performance among symbolization methods.

The evidence gathered in these simulations is represented in a series of heatmaps. Figure 1 presents a summary of the 24 comparisons we run in our exercises. Figures 2, 3, 4 and 5 depict more detailed heatmaps, corresponding to the linear, non-linear, random walk and continuous cases, respectively. Each of the latter figures shows the heatmap of the results with the original symbolic series and the heatmap for the differentiated series.

To generate the heatmaps, we consider for each possible relation among the variables, a value +1 if a the relation is detected at least at a 10% significance level for a symbolization method (including keeping the original numerical series) and for a causality detection procedure, and −1 if that combination yields an error or does not find a relation. The addition over all possible relations (Figure 1) or the specific relations (in the different simulation exercises) leads to a color scale in which the more positive results are depicted in red and the negative ones in blue.

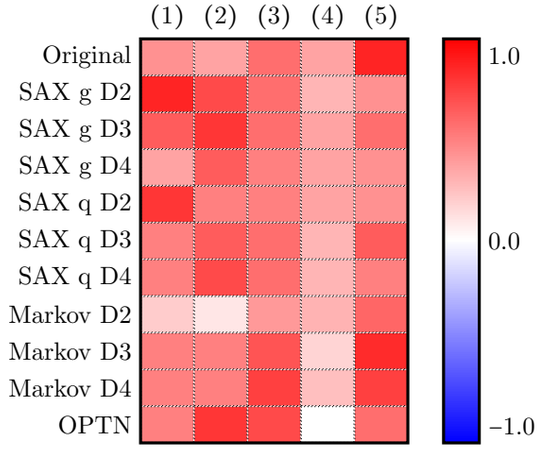
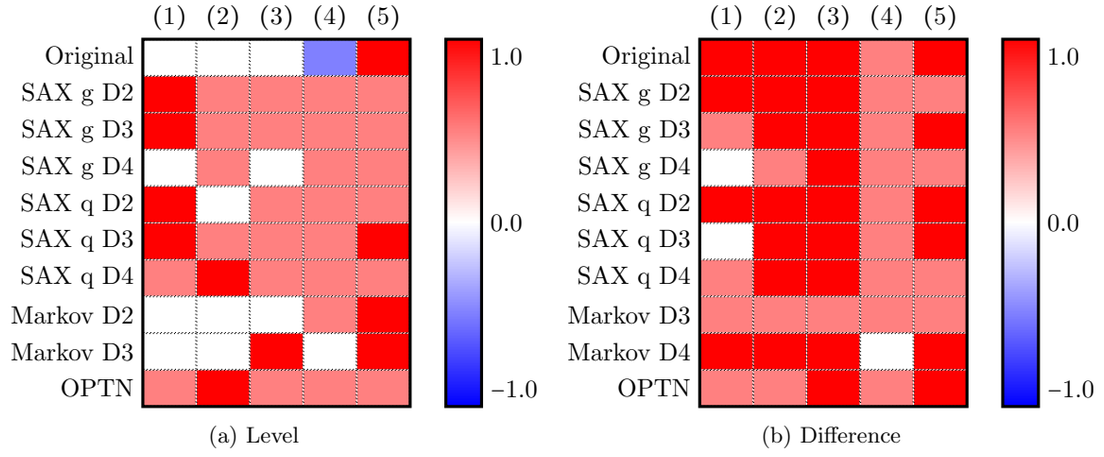


Figure 1: Full Heat Map

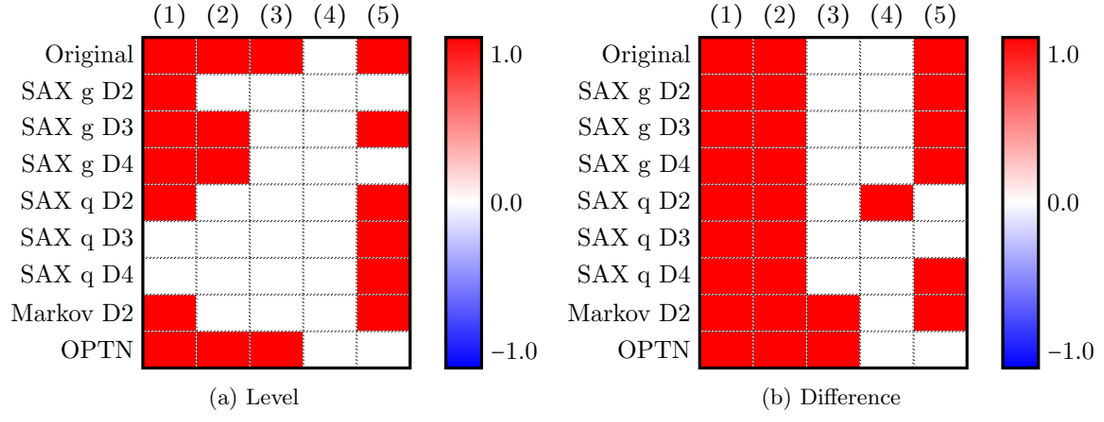
(1) is TE (Q = 0.5), (2) is TE (Q = 0.9), (3) is GRANGER (2 lags), (4) is PCMCI (lag= 0) and (5) is PCMCI (lag> 0).



(a) Level (b) Difference

Figure 2: Linear heat Map

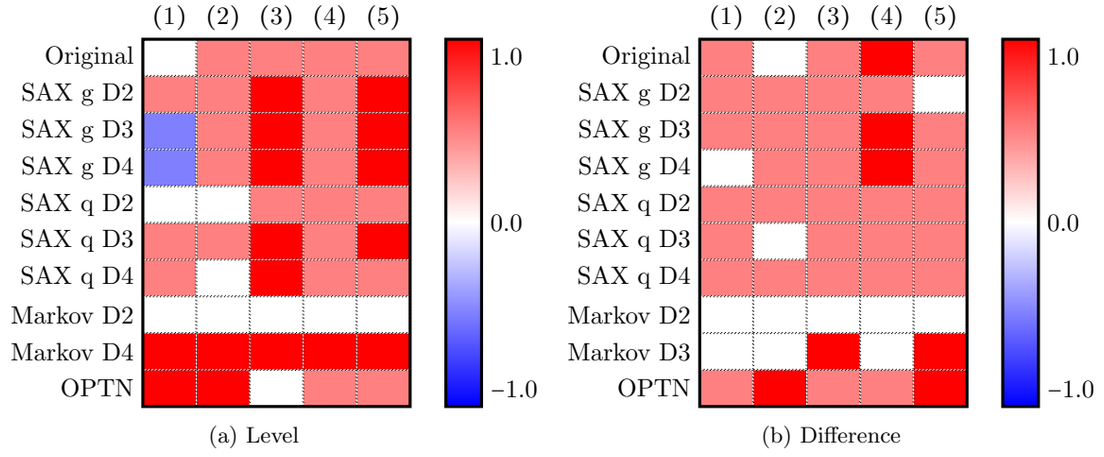
(1) is TE (Q = 0.5), (2) is TE (Q = 0.9), (3) is GRANGER (2 lags), (4) is PCMCI (lag= 0) and (5) is PCMCI (lag> 0).



(a) Level (b) Difference

Figure 3: Non-Linear heat Map

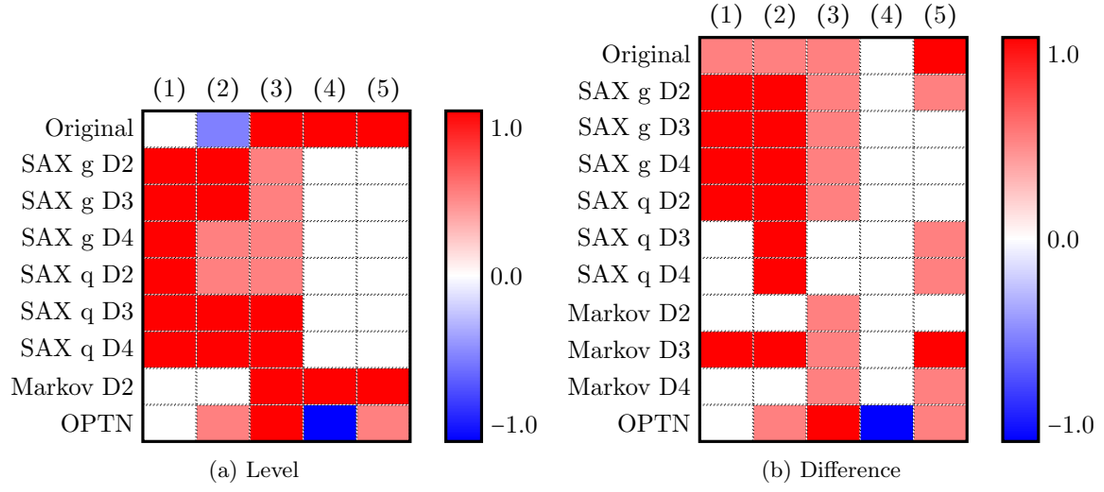
(1) is TE (Q = 0.5), (2) is TE (Q = 0.9), (3) is GRANGER (2 lags), (4) is PCMCI (lag= 0) and (5) is PCMCI (lag> 0).



(a) Level (b) Difference

Figure 4: Random-walk heat map

(1) is TE (Q = 0.5), (2) is TE (Q = 0.9), (3) is GRANGER (2 lags), (4) is PCMCI (lag= 0) and (5) is PCMCI (lag> 0).



(a) Level (b) Difference

Figure 5: Continuous dynamics heat map

(1) is TE (Q = 0.5), (2) is TE (Q = 0.9), (3) is GRANGER (2 lags), (4) is PCMCI (lag= 0) and (5) is PCMCI (lag> 0).

The main takeaways of these exercises are very general. For instance, we can recommend using differentiation when possible, since the symbolic series seem to perform better under differentiation. We also can recommend not to get attached a single symbolization methodology to capture the possible causal directions in the data. PCMCI yields the worst results with 0 lags, but seemingly the best ones with at least one lag. Similarly, Markow switching with more regimes seems to perform better.

4.6 Real world data: Google searches and the exchange rate in Argentina

In this exercise, we run an exercise that is frequent in economics and the social sciences in general. We focus on an ongoing debate about the *dollarization* of the economy in Argentina, namely the strong trend in both firms and households of saving in US dollars. Two contrary views are held in the public discourse

about this issue. One is that this is a precautionary measure in a highly inflationary economy, while the other is that it is just the result of an ingrained cultural trend. More information about this analysis can be seen in Delbianco et al. (2021).

We translate this question into one about the causal direction between four time series, between 2004 and 2019:

- Google searches ( ): searches about the value of the US dollar in the black market.
- Black market value ( ): the "true" (as evaluated by the public) exchange rate between the Argentinean peso and the US dollar.
- Consumer confidence (ICC): a consumer confidence index, measured by a private and independent institution in Argentina (CIF-UTDT).
- Inflation rate (π): the rate of variation of the consumer price index in Argentina.

The motivation for differencing, the pre-filtering steps, and the treatment of seasonality are discussed in Delbianco et al. (2021). The autocorrelation functions and the original time series of the four main variables in this study are presented in Figures 7 and 8 in the Appendix.

The cautionary interpretation indicates that the causality goes from the economic variables to the Google searches, while the "cultural" hypothesis means that people's online inquiries lead them to buy more dollars, increasing the value of the exchange rate, and doing so forces an increase in consumer prices, which in turn deteriorates the confidence in the economy Delbianco et al. (2021).

Since the underlying data generation model is unknown, we will just show the results of two symbolic transformations on the four series to apply two causality tests. Figure 6 depicts the statistically significant results of this exercise.

We can see that no causal relation can be detected unambiguously, since both the existence of a relation and its direction may vary with the number of symbols in the dictionary (2, 3 or 4), the method of transformation (SAX or Markov switching) or the causality test applied (transfer entropy or Granger).

5 Conclusion

We briefly presented the main tools of STSA, in particular focusing on the detection of causal relations among time series. The method of applying symbolic transformations to time series can help reduce the impact of noise and schematize the behavior of the series. But its combination with causality detection procedures may lead to the appearance of false causal relations or the failure to detect real ones.

The relative frequency of detections of false causal relationships compared to failed detections of real ones hints towards a potential prevalence of statistical artifacts generated by the combination of discretization with methods of causal detection developed with wider ranges of values in mind.

Of course, the exercises carried out here are just preliminary. We have only used three symbolization methods and three causality detection procedures. Other methods may provide stronger results, contributing to the development of new techniques. The goal would be to design "causality-preserving symbolization" procedures, or causality tests that may be tailored to the symbolization procedures. Both approaches would use the available information more efficiently.

Author's declarations:

Conflict of Interest Statement: The authors have no conflicts to disclose.
Data availability: The data that support the findings of this study are available from the corresponding

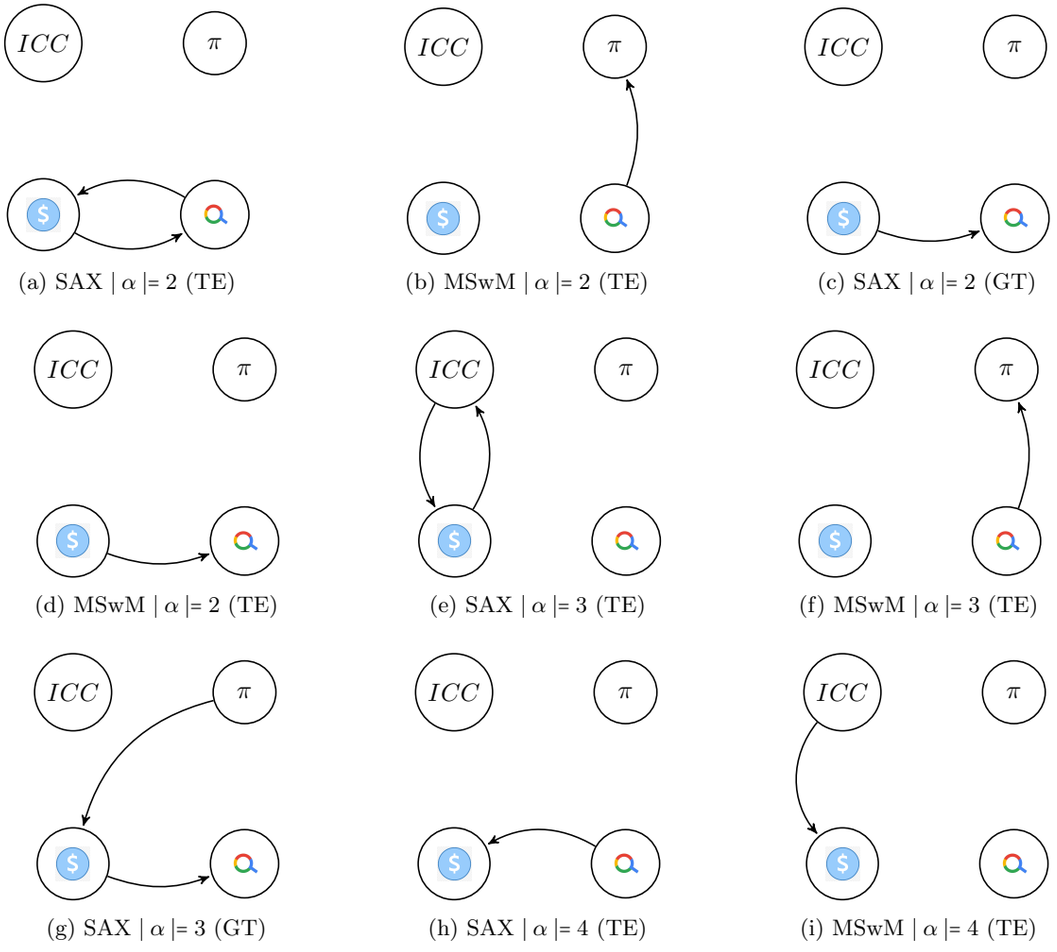


Figure 6: Detection of causal relations in symbolic specifications

The result of each test is represented graphically, indicating with directed edges the relations among variables. With SAX or MSwM (Markov Switching Model) we indicate the method of transformation applied, $|\alpha|$ indicates the number of symbols in the dictionary and in parentheses we indicate which test is applied, Transfer Entropy (TE) or Granger's Test (GT). We present only cases in which statistically significant results are obtained.

author upon reasonable request.

Author Contributions: All authors contributed equally.

Funding: This research was partially funded by the grant PGI-24-E172 UNS.

References

Bois, A. (2024). *Topological data analysis for time series*. PhD thesis, Université Paris-Saclay.

Brida, J. G. (2006). Multiple regimes model reconstruction using symbolic time series methods. *International Journal of Applied Mathematics & Statistics*, 5(S06):19–40.

Brida, J. G., Gómez, D. M., and Risso, W. A. (2009). Symbolic hierarchical analysis in currency markets: An application to contagion in currency crises. *Expert Systems with applications*, 36(4):7721–7728.

Brida, J. G. and Punzo, L. F. (2003). Symbolic time series analysis and dynamic regimes. *Structural Change and Economic Dynamics*, 14(2):159–183.

Chao, X. M. H. (2011). Analysis and forecasting of financial returns based on symbolic time series method. *Chinese Journal of Management Science*, page 05.

Daw, C. S., Finney, C. E. A., and Tracy, E. R. (2003). A review of symbolic analysis of experimental data. *Review of Scientific instruments*, 74(2):915–930.

Delbianco, F., Fioriti, A., Tohmé, F., and Contiggiani, F. (2021). A tale of two narratives: assessing the sociological hypothesis of the appeal of the us dollar in argentina. *Quality & Quantity*, pages 1–19.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: journal of the Econometric Society*, pages 424–438.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the econometric society*, pages 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regime. *Journal of econometrics*, 45(1-2):39–70.

Hamilton, J. D. (1994). *Time Series Analysis*, volume 2. Princeton university press Princeton, NJ.

Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., and Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *Physics Reports*, 441(1):1–46.

Lin, J., Keogh, E., Lonardi, S., and Chiu, B. (2003). A symbolic representation of time series, with implications for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, pages 2–11.

Lin, J., Keogh, E., Wei, L., and Lonardi, S. (2007). Experiencing sax: a novel symbolic representation of time series. *Data Mining and knowledge discovery*, 15(2):107–144.

Lonardi, J. and Patel, P. (2002). Finding motifs in time series. In *Proc. of the 2nd workshop on temporal data mining*, pages 53–68.

Malinowski, S., Guyet, T., Quiniou, R., and Tavenard, R. (2013). 1d-sax: A novel symbolic representation for time series. In *International Symposium on Intelligent Data Analysis*, pages 273–284. Springer.

Pourahmadi, M. (2001). *Foundations of Time Series Analysis and Prediction Theory*. John Wiley and Sons, NY.

Ruan, Y., Donner, R. V., Guan, S., and Zou, Y. (2019). Ordinal partition transition network based complexity measures for inferring coupling direction and delay from time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(4).

Runge, J., Nowack, P., Kretschmer, M., Flaxman, S., and Sejdinovic, D. (2019). Detecting and quantifying causal associations in large nonlinear time series datasets. *Science advances*, 5(11):eaau4996.

Sanchez-Espigares, J. A. and Lopez-Moreno, A. (2021). *MSwM: Fitting Markov Switching Models*. R package version 1.5.

Schreiber, T. (2000). Measuring information transfer. *Physical review letters*, 85(2):461.

Seversky, L. M., Davis, S., and Berger, M. (2016). On time-series topological data analysis: New data and opportunities. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 59–67.

Shahriari, Z., Algar, S. D., Walker, D. M., and Small, M. (2023). Ordinal poincaré sections: Reconstructing the first return map from an ordinal segmentation of time series. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 33(5).

Simon, B., Thomas, D., Franziska J., P., and David J., Z. (2019). Rtransferentropy — quantifying information flow between different time series using effective transfer entropy. *SoftwareX*, 10(100265):1–9.

Small, M. (2013). Complex networks from time series: Capturing dynamics. In *2013 IEEE International Symposium on Circuits and Systems (ISCAS)*, pages 2509–2512. IEEE.

Song, Q., Li, H., Huang, J., Huang, P., Tan, X., Tao, Y., Shi, C., and Zeng, G. (2023). Weak signal extraction in non-stationary channel with weak measurement. *Communications Physics*, 6(1):370.

Spirtes, P. (1991). Detecting causal relations in the presence of unmeasured variables. In *Uncertainty in Artificial Intelligence*, pages 392–397. Elsevier.

Vogt, W. P. and Johnson, R. B. (2015). *The SAGE dictionary of statistics & methodology: A nontechnical guide for the social sciences*. Sage publications.

Appendix

Table 1: $X_1 \longrightarrow Y_1$? ($Y_1 \longrightarrow X_1$?)

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (ns)	*** (ns)	*** (*)	ns (ns)	*** (ns)
SAX g D2	*** (ns)	*** (ns)	*** (***)	*** (***)	*** (***)
SAX g D3	*** (ns)	*** (ns)	*** (**)	ns (ns)	** (**)
SAX g D4	ns (ns)	*** (ns)	*** (**)	** (**)	*** (**)
SAX q D2	*** (ns)	*** (ns)	*** (***)	*** (***)	*** (**)
SAX q D3	** (ns)	*** (ns)	*** (**)	ns (ns)	*** (ns)
SAX q D4	ns (ns)	*** (ns)	*** (***)	** (**)	*** (**)
Marxov D2	** (ns)	*** (ns)	*** (ns)	ns (ns)	** (ns)
Marxov D3	ns (ns)	ns (ns)	*** (ns)	*** (***)	*** (ns)
OPTN	ns (ns)	** (ns)	*** (***)	*** (***)	*** (***)

Linear causal relation

Table 2: $X_2 \longrightarrow Y_1$**?** ($Y_1 \longrightarrow X_2$**?**)

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	** (ns)	** (*)	ns (***)	** (**)	ns (ns)
SAX g D2	ns (ns)	ns (**)	ns (ns)	ns (ns)	ns (ns)
SAX g D3	ns (ns)	ns (***)	ns (ns)	ns (ns)	ns (ns)
SAX g D4	** (ns)	*** (ns)	ns (**)	ns (ns)	ns (ns)
SAX q D2	ns (ns)	** (**)	ns (ns)	ns (ns)	ns (ns)
SAX q D3	ns (ns)	ns (***)	ns (ns)	ns (ns)	ns (ns)
SAX q D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
Marxov D2	** (*)	*** (*)	(**)	ns (ns)	ns (ns)
OPTN	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)

Linear causal relation

Table 3: $X_1 \longrightarrow Y_1$**?** ($Y_1 \longrightarrow X_1$**?): in differences**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	*** (ns)	*** (ns)	*** (ns)	ns (ns)	*** (ns)
SAX g D2	*** (ns)	*** (ns)	*** (ns)	ns (ns)	*** (ns)
SAX g D3	ns (ns)	*** (ns)	*** (ns)	ns (ns)	*** (ns)
SAX g D4	ns (*)	** (*)	*** (ns)	ns (ns)	*** (ns)
SAX q D2	*** (ns)	*** (ns)	*** (ns)	ns (ns)	*** (ns)
SAX q D3	ns (*)	*** (ns)	*** (ns)	ns (ns)	*** (ns)
SAX q D4	ns (ns)	* (ns)	*** (ns)	ns (ns)	*** (ns)
Marxov D3	ns (ns)	ns (ns)	*** (**)	ns (ns)	*** (***)
Marxov D4	** (ns)	** (ns)	*** (ns)	ns (ns)	*** (ns)
OPTN	ns (ns)	ns (ns)	*** (ns)	ns (ns)	*** (ns)

Linear causal relation

Table 4: $X_2 \longrightarrow Y_1$**?** ($Y_1 \longrightarrow X_2$**?): in differences**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	** (ns)
SAX g D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (**)
SAX q D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (**)
Marxov D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
OPTN	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)

Linear causal relation

Table 5: $X_1 \longrightarrow Y_1'$**?** ($Y_1' \longrightarrow X_1$)**?**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	*** (NS)	*** (ns)	*** (NS)	** (**)	*** (ns)
SAX g D2	*** (ns)	*** (**)	*** (**)	*** (***)	*** (**)
SAX g D3	*** (ns)	*** (ns)	*** (*)	ns (ns)	*** (ns)
SAX g D4	*** (ns)	*** (ns)	*** (***)	ns (ns)	** (**)
SAX q D2	*** (ns)	*** (**)	*** (**)	ns (ns)	*** (ns)
SAX q D3	*** (*)	*** (***)	*** (**)	ns (ns)	*** (ns)
SAX q D4	*** (*)	*** (***)	*** (**)	ns (ns)	*** (ns)
Marxov D2	*** (ns)	*** (*)	*** (*)	ns (ns)	** (ns)
OPTN	** (ns)	*** (ns)	*** (ns)	ns (ns)	*** (***)

Non-Linear causal relation

Table 6: $X_1 \longrightarrow Y_1'$**?** $(Y_1' \longrightarrow X_1$**?): in differences**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	*** (ns)	*** (ns)	*** (*)	ns (ns)	*** (ns)
SAX g D2	*** (ns)	*** (ns)	*** (*)	ns (ns)	*** (ns)
SAX g D3	*** (ns)	*** (ns)	*** (**)	ns (ns)	*** (ns)
SAX g D4	*** (ns)	*** (ns)	*** (***)	** (**)	*** (ns)
SAX q D2	*** (ns)	*** (ns)	*** (*)	*** (ns)	ns (ns)
SAX q D3	*** (ns)	*** (ns)	*** (**)	*** (**)	*** (**)
SAX q D4	** (ns)	*** (ns)	*** (***)	ns (ns)	*** (ns)
Marxov D2	*** (ns)	*** (ns)	*** (ns)	*** (***)	*** (ns)
OPTN	** (ns)	*** (ns)	*** (ns)	*** (***)	*** (***)

Non-Linear causal relation

Table 7: $X_1 \longrightarrow Y_1''$**?** $(Y_1'' \longrightarrow X_1$**?)**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (ns)	** (ns)	*** (***)	*** (***)	*** (***)
SAX g D2	ns (ns)	ns (ns)	** (ns)	ns (ns)	** (ns)
SAX g D3	ns (ns)	ns (ns)	** (ns)	ns (ns)	** (ns)
SAX g D4	ns (ns)	ns (ns)	*** (ns)	ns (ns)	** (ns)
SAX q D2	* (**)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D3	ns (ns)	ns (ns)	** (ns)	ns (ns)	** (ns)
SAX q D4	** (ns)	ns (ns)	** (ns)	ns (ns)	ns (ns)
Marxov D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
OPTN	*** (ns)	* (ns)	*** (**)	ns (ns)	*** (ns)

Random walk

Table 8: $X_2 \longrightarrow Y_1''$**?** $(Y_1'' \longrightarrow X_2$**?)**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	* (ns)	ns (**)	ns (ns)	ns (ns)	ns (ns)
SAX g D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D3	* (*)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D4	* (*)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D2	*** (ns)	*** (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D4	* (ns)	* (ns)	ns (ns)	ns (ns)	ns (ns)
Marxov D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
OPTN	ns (ns)	ns (ns)	ns (***)	ns (ns)	ns (***)

Random walk

Table 9: $X_1 \longrightarrow Y_1''$**?** $(Y_1'' \longrightarrow X_1$**?): in differences)**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (ns)	ns (*)	*** (***)	** (ns)	*** (***)
SAX g D2	*** (**)	*** (**)	*** (***)	** (**)	*** (***)
SAX g D3	*** (***)	*** (**)	*** (***)	*** (ns)	*** (***)
SAX g D4	ns (**)	*** (*)	*** (***)	** (ns)	*** (***)
SAX q D2	*** (*)	*** (**)	*** (***)	ns (ns)	*** (***)
SAX q D3	*** (*)	*** (**)	*** (***)	ns (ns)	*** (***)
SAX q D4	** (***)	*** (**)	*** (***)	ns (ns)	*** (***)
Marxov D2	ns (ns)	* (***)	*** (***)	*** (***)	*** (***)
Marxov D3	ns (ns)	ns (ns)	*** (ns)	ns (ns)	*** (ns)
OPTN	ns (ns)	*** (ns)	*** (***)	*** (***)	*** (ns)

Random walk

Table 10: $X_2 \longrightarrow Y_1''$**?** ($Y_1'' \longrightarrow X_2$**?): in differences**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	** (ns)
SAX g D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D3	ns (ns)	ns (**)	ns (ns)	ns (ns)	ns (ns)
SAX q D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
OPTN	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)

Random walk

Table 11: $X_1 \longrightarrow Y_1'''$**?** ($Y_1''' \longrightarrow X_1$**?**)

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (ns)	ns (ns)	*** (***)	*** (***)	*** (***)
SAX g D2	*** (**)	*** (**)	*** (ns)	ns (ns)	ns (ns)
SAX g D3	*** (**)	*** (**)	ns (***)	ns (ns)	ns (ns)
SAX g D4	*** (**)	*** (ns)	*** (ns)	ns (ns)	ns (ns)
SAX q D2	*** (**)	ns (***)	ns (***)	ns (ns)	ns (ns)
SAX q D3	*** (**)	* (***)	* (***)	ns (ns)	ns (ns)
SAX q D4	** (***)	* (***)	* (***)	ns (ns)	ns (ns)
Marxov D2	*** (***)	*** (***)	*** (***)	*** (***)	*** (***)
OPTN	ns (ns)	*** (ns)	*** (***)	ns (ns)	** (ns)

A continuous dynamics

Table 12: $X_2 \longrightarrow Y_1'''$**?** ($Y_1''' \longrightarrow X_2$**?**)

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (ns)	** (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX g D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
SAX q D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	ns (ns)
Marxov D2	*** (***)	*** (***)	ns (ns)	ns (ns)	ns (ns)
OPTN	ns (ns)	ns (ns)	ns (ns)	** (**)	ns (ns)

A continuous dynamics

Table 13: $X_1 \longrightarrow Y_1'''$**?** ($Y_1''' \longrightarrow X_1$**?): in differences**

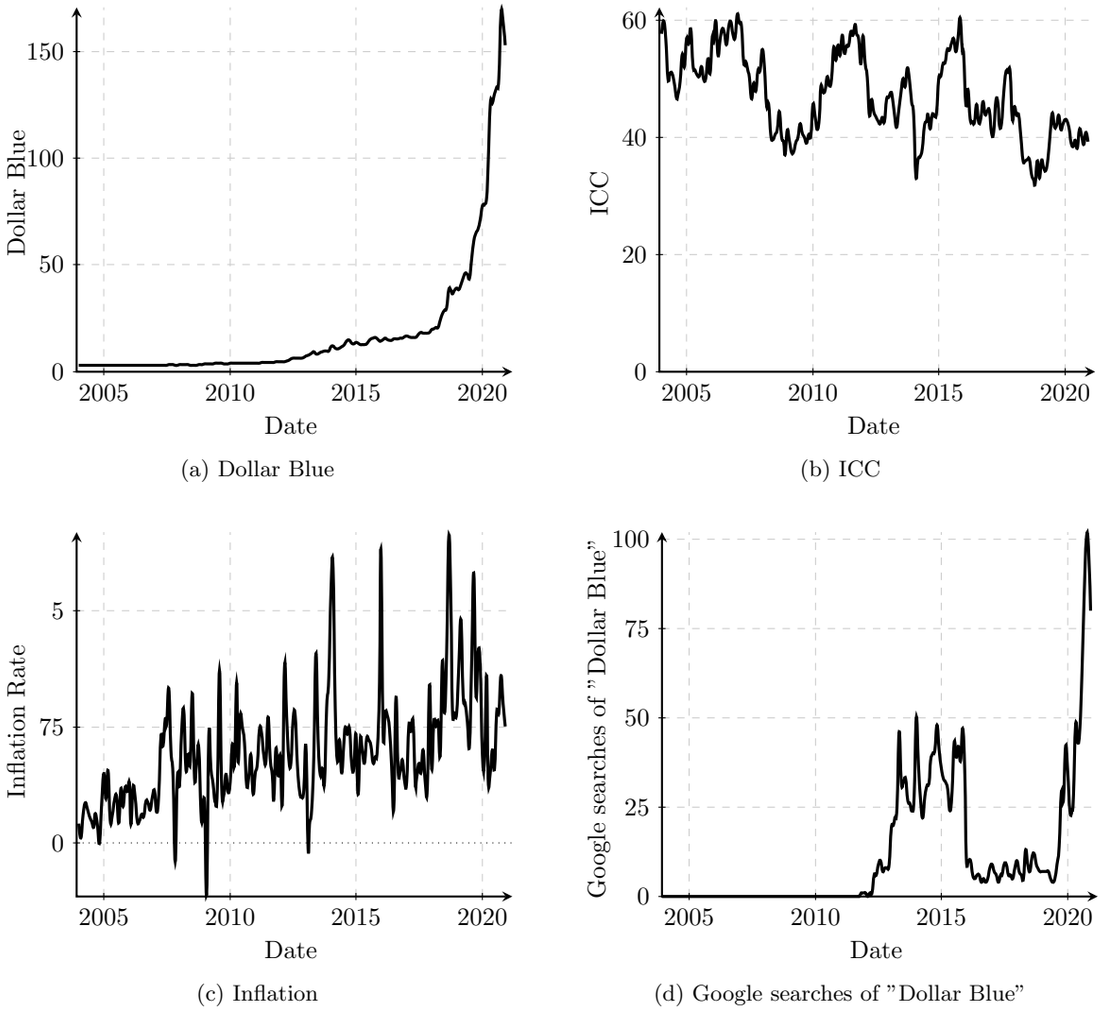
Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
Original	ns (***)	ns (***)	*** (***)	ns (ns)	*** (***)
SAX g D2	*** (***)	*** (***)	*** (ns)	ns (ns)	ns (***)
SAX g D3	*** (***)	*** (***)	*** (ns)	ns (ns)	ns (***)
SAX g D4	*** (***)	*** (***)	*** (ns)	ns (ns)	ns (***)
SAX q D2	*** (***)	*** (***)	*** (ns)	ns (ns)	ns (***)
SAX q D3	ns (ns)	*** (**)	*** (ns)	ns (ns)	ns (***)
SAX q D4	ns (ns)	*** (**)	ns (***)	ns (ns)	ns (***)
Marxov D2	*** (**)	*** (***)	*** (ns)	ns (ns)	ns (***)
Marxov D3	*** (***)	*** (***)	*** (ns)	ns (ns)	*** (***)
Marxov D4	*** (***)	*** (***)	*** (ns)	ns (ns)	ns (***)
OPTN	ns (ns)	*** (ns)	*** (***)	ns (ns)	ns (**)

A continuous dynamics

Table 14: $X_2 \longrightarrow Y_1'''$**?** $(Y_1''' \longrightarrow X_2$**?): in differences**

Symbolic series	TE (Q = 0.5)	TE (Q = 0.9)	GRANGER (2 lags)	PCMCI (lag= 0)	PCMCI (lag> 0)
SAX g D2	ns (ns)	ns (ns)	ns (***)	ns (ns)	ns (ns)
SAX g D3	ns (ns)	ns (ns)	ns (ns)	ns (ns)	** (ns)
SAX g D4	ns (ns)	ns (ns)	ns (ns)	ns (ns)	** (ns)
SAX q D2	ns (ns)	ns (ns)	ns (ns)	ns (ns)	** (ns)
SAX q D3	ns (ns)	ns (ns)	* (ns)	ns (ns)	ns (ns)
SAX q D4	ns (ns)	ns (ns)	* (ns)	ns (ns)	ns (ns)
Marxov D2	*** (ns)	*** (ns)	ns (ns)	ns (ns)	** (ns)
Marxov D3	ns (***)	ns (***)	ns (ns)	ns (ns)	ns (ns)
Marxov D4	*** (ns)	*** (ns)	ns (ns)	ns (ns)	ns (ns)
OPTN	ns (***)	ns (***)	ns (ns)	** (ns)	ns (ns)

A continuous dynamics



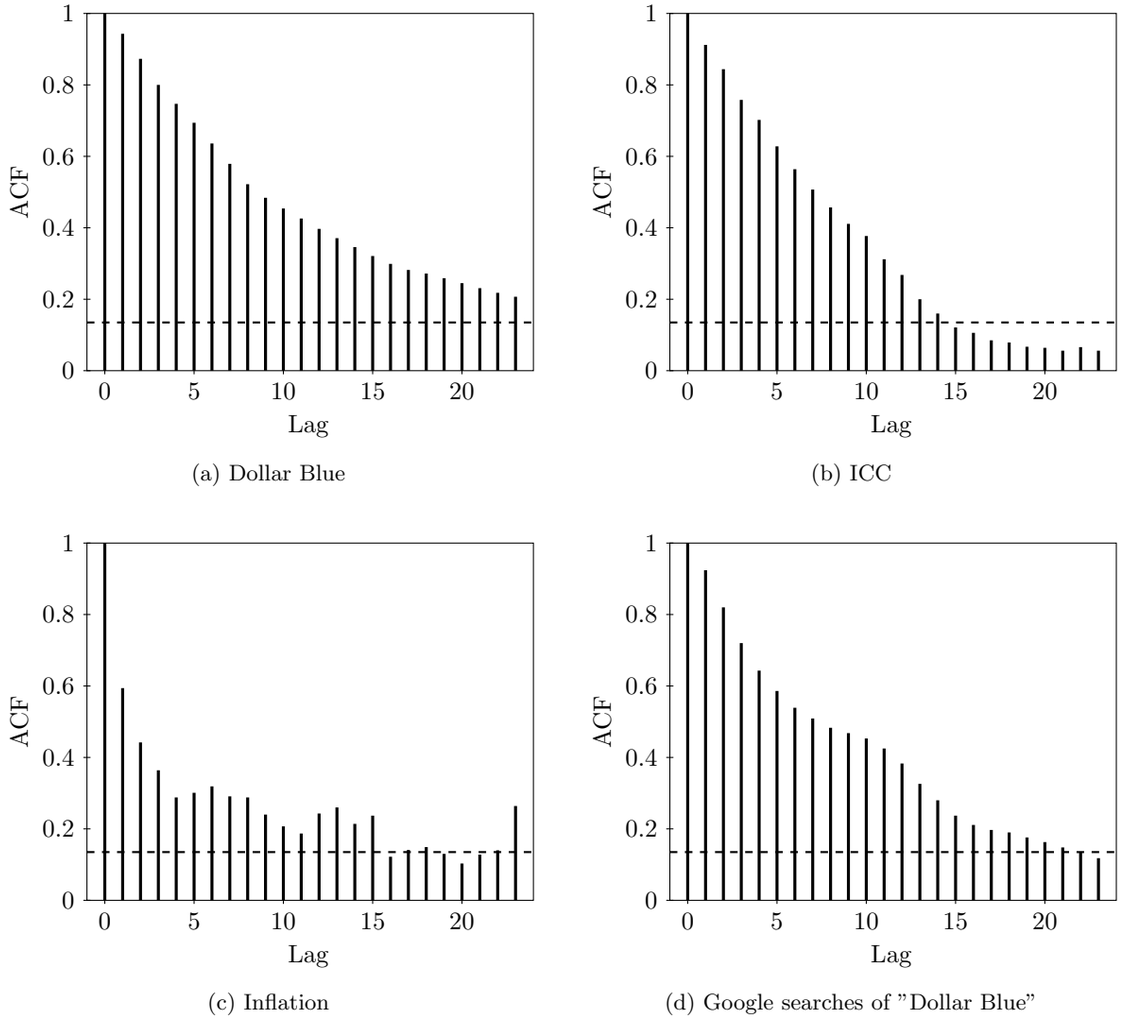
(a) Dollar Blue

(b) ICC

(c) Inflation

(d) Google searches of "Dollar Blue"

Figure 7: Time series plot of the empirical application



(a) Dollar Blue

(b) ICC

(c) Inflation

(d) Google searches of "Dollar Blue"

Figure 8: Autocorrelation plot of the empirical series used in the empirical application