CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.

CIDE

DETERMINING CAUSAL INFERENCE IN LINEAR AND NON-LINEAR TIME-SERIES
USING CONVERGENT CROSS MAPPING
AN APPLICATION OF GOVERNMENT EXPENDITURE AND ECONOMIC GROWTH RELATION
IN MEXICO 1980 - 2015

TESINA

QUE PARA OBTENER EL GRADO DE

MAESTRA EN ECONOMÍA

PRESENTA

RUBI TONANTZIN GUTIÉRREZ VILLANUEVA

DIRECTOR DE LA TESINA:

DR. DANIEL VENTOSA - SANTAULÀRIA

CIUDAD DE MÉXICO.   JUNIO, 2016

Education is our passport to the future,
for tomorrow belongs to the people who
prepare for it today.

Malcom X

ACKNOWLEDGEMENTS

Abstract

This study has two major purposes: (1) to identify causal inference in time-series using *Granger causality tests* and *Convergent Cross Mapping (Sugihara et al., 2012)* (2) to investigate the economic growth and government expenditure relation in Mexico. *Convergent Cross Mapping (CCM)* has shown a high potential to perform causal inference in complex systems and non-linear system and has been used as an alternative approach to Granger causality. One one hand, we show that *CCM* fails to infer causality direction in linear time-series and in time-series with structural breaks. On the other hand, we demonstrate that *Toda-Yamamoto test (Toda and Yamamoto, 1995)* successfully detects causal relation in linear systems and systems with structural breaks. Besides, we evaluate the causal relation between government expenditure and economic growth in Mexico then we evaluate the validity of Wagner's law and the Keynesian view. The empirical results suggests that Wagner's law holds for Mexico for the period 1980 to 2010.

Keywords: *Causality*, *Convergent Cross Mapping*, *Granger Causality Tests*, *Time Series Analysis*, *Government Expenditure*, *Economic Growth*, *Mexico*

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Causality plays a fundamental and controversial role in the sciences. The objective of research in causality is to build models that try to explain causal inference and causal explanation. For this reason, different disciplines, from epidemiology to biology, from econometrics to physics, have made use of probabilistic and statistics methods in order to infer causal relations. And as a consequence philosophers, scientists, physics, mathematicians, economists, computer scientists and many others have studied the field of causation. The exploration of causation is a knowledge field, which started with the ancient Greeks three thousand years ago. In Metaphysics, Aristotle maintains that knowing is knowing attending to the causes (Agueda, 2011).

> *But yet we think that knowledge and understanding belong to art rather than to experience, and we suppose artists to be wiser than men of experience (which implies that Wisdom depends in all cases rather on knowledge); and this because the former know the cause, but the latter do not. For men of experience know that the thing is so, but do not know why, while the others know the **why** and the cause.*

Causation plays a different role when examined from different fields. As a social science, economics was conceived as early as a science of causes. For example, causal inferences and causal motivations in economics are motivated by policy questions. In this context, the goals of policy evaluation are to consider the impact of policy interventions on the economy, to estimate their consequences for economic welfare, and to forecast the effects of new policies.

The base of economic causality was proposed by the philosopher–economists David Hume contributed to the philosophy of causation in his books *A Treatise of Human Nature* and *An Enquiry concerning Human Understanding*. Both books start with Hume's axiom known as the *Copy Principle* which states that all components of our thoughts come from experience. Also he argued that our idea of necessary connection, which he concedes as a characteristic element of causality can appear from our experience.

J. S. Mill another philosopher/economist was skeptical about causality application to economics. In 1984, Mill published his textbook *Principles of Political Economy*. In this work, Mill argued that economics was an 'inexact and separate science', whose principles were essentially known a priori and which held only subject to ceteris paribus clauses.

The 18th century, developed the conceptions of causality as *" we may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second"*. Then, economics was deeply affected by the philosophy of causal determinism, which the natural sciences embraced throughout the nineteenth century. That philosophy is most famously espoused by the philosopher Pierre-Simon Laplace thus (Morck and Yeung, 2011):

> *We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes*

For this intellect, dubbed **Laplace's demon**, every event is a cog in a mechanical chain stretching back to the beginning of the universe.

A century after Hume's definition, Granger proposed the idea of *Granger causality* in 1969. In that paper, the author described the **causal relationships** between variables in economic models. This idea is simple one but has a key requirement which is named separability (information about a causative factor is independently to that variable and can be taken out by eliminating the variable from the model) and when it is violated Granger causality calculations are no longer valid.

Trying to solve last problem, Sugihara et al. (2012) introduced a technique called convergent cross mapping (CCM) for computing correlation between variables, based on non-linear state space reconstruction that can distinguish causality from correlation. CCM is fast becoming a key instrument in wide range of fields including neuroscience, biomedical, climatology and social interaction. Its wide range of applications already appearing for the CCM method is a proof of its importance in time-series causality research. In order to know the power of this new technique, we present an extensive set of examples where the performance of Granger causality and CCM procedure is measured and compared. These processes will have different natures as

stochastic, non-linear or causal.

Besides, we study the relationship between government expenditure and economic growth in Mexico for the period 1980 - 2015. Current research on this relationship is focused on how developing countries have faced structural adjustment and enhanced their revenue situation. We focus on the adjustment process of the public sector. If a government's total expenditures exceeds its revenue a high fiscal deficit results and as a consequence, economic performance could be affected. For this reason the relation between growth of public expenditure and Gross Domestic Product has considerable attention. For this reason, CCM method is used to estimate the long-run causality of economic indicators for Mexico. This new methodology will be compared with existing causal testing procedures.

The rest of this thesis is organized as follows. We begin with a review of previous studies based on Sugihara et al. (2012) and government expenditure and economic growth relation. Next chapter introduces a basic idea an algorithm of CCM. Empirical results will presented in section four. Last part of the thesis describes how economic growth interacts with government expenditure for Mexico.

Chapter 2

Literature Review

In economics, the most influential approach to causality was developed by Clive W. J. Granger (Granger, 1969) named **Granger causality** in his honor. In his article, Granger developed probabilistic approaches to causality, those approaches are calculated to identify cause effects with a factor that increases the probability of the effect. The initial definition of causality is a general concept based on minimum variance in a model with a completed set of variables against a model with a restricted set of those variables. This original definition of causality had not been restricted in order to reach a form which can be tested, thus Granger (1969) used the predictability of two variables to define causality between them. Then Granger (1969) defined causality when a variable x_t causes y_t if x_t contains past information that helps to predict information of variable y_t in present time.

Next sections explain with more detail **Granger causality** and its weakness with non-stationary variables and deterministic models.

2.1 Granger Causality test

The method developed by Granger (1969) is applied to stationary variables, if this method allows non-stationary time-series then variance will depend on time t, as a consequence, causality may be altered over time.

The first definition given by Granger (1969) supposed a set of variables u_t with all information in the universe accumulated from $t-1$ and $u_t - y_t$ which denoted all this information less information contained in y_t. Granger (1969) defined causality as a variable y_t **causes** variable x_t if exist a model with all variables in the universe u_t fits better than the model with all variables less y_t information. The best fitted model will be the model with minimum variance, in this case y_t **is causing** x_t if the variance in the model with y_t is less than the model without y_t.

$$\sigma^2(x_t \mid \bar{u}_t) < \sigma^2(x_t \mid \overline{u_t - y_t}), \tag{2.1}$$

Granger (1969) argued that the variance is not the proper criterion to use to measure good predictors for x_t. This means that if some other criteria were used to predict variables then it is possible to reach different conclusions about whether one variable is causing another. Although, the variance seem to be a natural criterion to use in connection with linear predictor, this criterion of causality is restricted in order to reach a criteria which can be tested.

With the last idea in mind, Granger (1969) used a definition of causality based entirely on the predictability of time-series. In this way, Granger causality can be defined as follows: "a time-series x_t is said **Granger causes** a time-series y_t, if past values of x_t contain information

that help predict y_t". One of the most used ways to test Granger causality is to model an autoregressive process of a bivariate vector autoregression (*VAR*). Given a finite lag m, the following equation is estimated

$$x_t = c_1 + \sum_{i=1}^{m} \alpha_i x_{t-i} + \sum_{i=1}^{m} \beta_i y_{t-i} + \epsilon_t, \qquad (2.2)$$

$$y_t = c_2 + \sum_{i=1}^{m} \gamma_i x_{t-i} + \sum_{i=1}^{m} \delta_i y_{t-i} + \eta_t,$$

where α, β, γ, δ, c_1 and c_2 are coefficients and ϵ_t and η_t are two uncorrelated white-noise time-series. The null hypothesis that will be tested is that y_t ***does not*** *Granger-cause* x_t, in other words

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_m = 0,$$

Thus, given Equation 2.2 y_t causes x_t if some β_i is not zero. Similarity, x_t ***Granger-cause*** y_t if some γ_i is not zero. Granger (1969) defined *feedback* relation if both of these events occur. Wald test could be used to test the null hypothesis in this cases.

It is important to emphasise that the definitions above assumed that only stationary time-series are involved. The existence of unit root in the variables and cointegration between variables could make the asymptotic inference invalid. As a consequence of lack of stationary in variables, we could fail to reject the null hypothesis when stationary variables must reject the hypothesis. Besides, it is possible to obtain a spurious regression, as a consequence, the regression model could have a high R^2 even when the time-series are independent of each other (Granger and Newbold, 1974). Before proceeding to examine unit root tests, it will be necessary

to define stationarity of time-series.

2.1.1 Stationarity of time-series

Stationarity is a property of a process which ensure that its mean, variance and other moments do not change when time is shifted. For example, white noise is a stationary time-series because it does not follow any trend and its mean and variance do not change over time. Stationarity can usually be defined as strict and weakly stationarity.

Definition 2.1.1. Strongly stationarity: The process y_t is strong stationary if the joint distribution $[y_t, y_{t+1}, \ldots, y_{t+k}, y_{t+k+1}]$ is equal for any t.

In other words, strict stationarity means that the joint distribution only depends on the *difference* k, not the time t.

Definition 2.1.2. Weakly stationarity: The process y_t is weakly stationary, or covariance-stationary if :

- $E(y_t) = E(y_{t-j}) = \mu,$

- $E[(y_t - \mu)^2] = E[(y_{t-j} - \mu)^2] = \sigma^2 < \infty,$

- $E[(y_t - \mu)(y_{t-j} - \mu)] = E[(y_{t-k} - \mu)(y_{t-j-k} - \mu)] = \gamma_j,$

where μ, σ^2 and γ_j are all constants.

In other words, a weak stationary time-series must have three characteristics: finite variance, constant mean and the second moment γ_j only depend on $(t - j)$.

If a time-series shows a trend or is affected by a persistent innovation, it is possible to test whether it is stationarity to get round the problem or to understand its possible effects. This kind of tests are called **unit root tests** and it can be used to determine if trending data should be first differenced or regressed on deterministic functions of time to convert the data stationary. The pioneering work on testing for a unit root in time-series was done by Dickey and Fuller (1979) in next section we explain the basic objective of the unit-root test.

2.1.2   Unit root test

Many economic time-series exhibit trending behaviour or non-stationary in the mean. Unit root tests can be used to determinate if trending data should be first differenced to obtain a stationary process. Furthermore unit root tests can be used to establish an order of integration. Order of integration, denoted $I(d)$ display the minimum number of differences required to obtain a stationary time-series.

To understand the notion of unit root, consider the following processes y_{t_1} , y_{t_2} and y_{t_3} define as:

$$y_{t_1} = u_1 + y_{t_1-1} + \mu_{t_1}, \tag{2.3}$$

$$y_{t_2} = u_2 + \beta t + \mu_{t_2}, \tag{2.4}$$

$$y_{t_3} = y_{t_3-1} + \mu_{t_3}, \tag{2.5}$$

Rewriting equation 2.3 as:

$$(1 - L)y_{t_1} = u_1 + \mu_{t_1}, \tag{2.6}$$

where L is the lag operator. In this case equation 2.6 has two roots, one of them is equal to one and as consequence the process has a unit root. Using equations 2.3, 2.4 and 2.5, we define equation 2.7

$$y_t = u + \beta t + y_{t-1} + \mu_t. \tag{2.7}$$

If we take 2.7, subtract y_{t-1} from both sides and introduce the artificial parameter γ, we obtain:

$$
\begin{aligned}
y_t - y_{t-1} &= \gamma u + \gamma \beta t + (\gamma - 1)y_{t-1} + \mu_t \\
&= a_0 + a_1 t + (\gamma - 1)y_{t-1} + \mu_t
\end{aligned} \tag{2.8}
$$

where by hypothesis $\gamma = 1$. Equation 2.8 supplies the basis for unit root test where if $(\gamma - 1) = 0$ gives confirmation of a random walk with a drift as equation 2.3, if $(\gamma - 1) < 0$ then the evidence favors the trend stationary. So, we have to test the next null hypothesis:

$$H_0 : \gamma = 1 \text{ or } (\gamma - 1) = 0 \Longrightarrow \text{ series contains a unit root}$$

$$H_1 : \gamma < 1 \text{ or } (\gamma - 1) < 0 \Longrightarrow \text{ series is stationaty}$$

Moreover, if the process is non-stationary and the first difference of the process is stationary, the process contains a unit root. The commonly used methods to test for the presence of unit

root are the Augmented Dickey-Fuller (ADF) tests (Dickey and Fuller, 1979), Phillip-Perron test (Phillips and Perron, 1988) and KPSS (Kwiatkowski et al., 1992).

In 1989, Perron discussed that most macroeconomic time-series are not characterized by a unit root but rather that persistence arises only from large and unusual shocks, and that the economy returns to deterministic trend after small and frequent shocks. Then Perron (1989) argued that if a process experiences a structural break then the unit root test is biased towards the no rejection of the null hypothesis. In fact, Perron used a modified Dickey-Fuller unit root test that includes dummy variables to account for one exogenous structural break. In 2005, Kapetanios developed a test for the unit root hypothesis against the emergence of an unspecified number of breaks in a time-series.

2.1.3 Structural breaks and Kapetanios test

Taking account the test of Perron, recent analysis of time-series data has been focused on more than one structural break. Kapetanios (2005) provided a test for the unit root hypothesis with drift but no breaks against an alternative hypothesis of a trend stationary process with an unspecified number of breaks in the trend or constant. The proposed test follows a sequential Dickey-Fuller test (explained above 2.1.2 and is defined as:

$$y_t = \mu_0 + \mu_1 t + \alpha y_{t-1} + \sum_{i=1}^{k} \gamma_i \Delta y_{t-i} + \sum_{i=1}^{m} \phi_i DU_{i,t} + \sum_{i=1}^{m} \varphi_i DT_{i,t} + \epsilon_t, \qquad (2.9)$$

$1 - \gamma(L)$ has all its roots outside the unit circle, where $\gamma(L) = \gamma_1 L + \ldots + \gamma_k L^k$. $DU_{i,t}$ and $DT_{i,t}$ are intercept and trend break dummy variables respectively and are defined by:

$$DU_{i,t} = 1(t > T_{b,i}),$$

$$DT_{i,t} = 1(t > T_{b,i})(t - T_{b,i}),$$

where $T_{b,i} + 1$ denotes the date of the ith structural break and $1(\cdot)$ is the indicator function.

The possibly spurious relationship between two variables has been solved with first difference each series and then it is possible to calculate the regression. However, a major problem with this kind of application is that *valuable long-run information may be lost*. Then another potential problem is to find a way to work with two possibly non-stationary time-series in a model that allows us to capture both short run and long run effects. The solution is called cointegration and it is the link between processes with the same order of integration and steady state equilibrium.

2.1.4 Cointegration

When there is evidence favorable to the unit root hypothesis in all variables is possible to develop models that conduct to stationary relations among the variables and where standard inference is possible. The criteria for stationarity among non-stationary variables is called cointegration.

Consider the simplest stochastic trend model:

$$y_t = \mu_y + \beta_y x_t + \epsilon_{yt}, \tag{2.10}$$

$$x_t = x_{t-1} + \epsilon_{xt}, \tag{2.11}$$

$$= \sum_{i=1}^{t} \epsilon_{xi},$$

where ϵ_{yt} and ϵ_{xt} are stationary series. In this case y_t and x_t contain a unit root. Besides, we can write y_t as:

$$y_t = \mu_y + \beta_y x_t + \epsilon_{yt}, \tag{2.12}$$

$$= \mu_y + \beta_y \{x_{t-1} + \epsilon_{xt}\} + \epsilon_{yt},$$

$$= \mu_y + \beta_y \sum_{i=1}^{t} \epsilon_{xi} + \epsilon_{yi},$$

now it is easy to see that y_t is a non-stationary process. In this model y_t and x_t are non-stationary time-series data with the same integration order. There are several tests of cointegration, the first test of cointegration was proposed by Engle and Granger (1987). The test captures equation 2.12 and sorts it as:

$$y_t - \mu_t - \beta_y x_t = \epsilon_{yt}, \tag{2.13}$$

The idea of Engle and Granger cointegration is to take a linear combination of variables with the same order of integration and obtain a variable with a less order of integration. In this case, ϵ_{yt} is stationary. Even, tests for cointegration (Engle and Granger, 1987) and cointegration rank have been developed as Johansen (1988) who based on VAR representation of time-series, simulation experiments have showed that the test is sensitive to values of parameters in finite samples and hence not very reliable for sample sizes that are typical for economic time-series.

Besides, Engel and Granger noted that while two variables are non-stationary and cointegrated, the standard Granger causal inference will be invalid.

To mitigate these problems, Toda and Yamamoto (1995) presented a simple way to overcome the problems in the testing that we find when VAR processes may have some unit roots. Toda-Yamamoto test is applicable whether the VAR's may be stationary (around a deterministic trend), integrated of an arbitrary order, or cointegrated of an arbitrary order. The following section will explain Toda-Yamamoto work.

2.1.5 Toda and Yamamoto test

Toda and Yamamoto (1995) based on VAR modeling, introduced a modified Wald test. Firstly, Toda-Yamamoto test find the maximum order of integration of the time-series and secondly an optimal number of lags. Toda and Yamamoto (1995) augmented Granger causality test procedure, which is based on the following assumptions.

Let y_t sequence be generated by the following function:

$$y_t = \beta_0 + \beta_1 t + \ldots + \beta_q t^q + \eta_t,$$

where $\eta_t e$ contains an order of integration equal to d. In particular, η_t is a vector autoregression with k lag length and it can be presented as :

$$\eta_t = J_1 \eta_{t-1} + \ldots + J_k \eta_{t-k} + \epsilon_t,$$

where k is assumed to be known and ϵ_t is an i.i.d sequence of random vectors with mean zero and finite variance.

Substituting $\eta_t = y_t - \beta_0 - \beta_1 t - \ldots - \beta_q t^q$ into first equation, getting

$$y_t = \gamma_0 + \gamma_1 t + \ldots + \gamma_q t^q + J_1 y_{t-1} + \ldots + J_k y_{t-k} + \epsilon_t,$$

Where γ_i is function of β_i. As order of integration $d > 0$, the order of trend γ might be lower than order q. Assume $d = 1$ and $q = 1$, $\gamma_2 = \gamma_3 = \ldots = \gamma_q = 0$. Then

$$y_t = \gamma_0 + \gamma_1 t + J_1 y_{t-1} + \ldots + J_k y_{t-k} + \epsilon_t,$$

Toda and Yamamoto (1995) were interested in the significance of coefficient of lagged y_t and not in whether the process y_t is integrated, cointegrated or stationary. Then the hypothesis is formulated as:

$$H_0 : J_1 = J_2 = \ldots = J_k = 0$$

Then k is the optimal lagged length and any additional lag are indifferent from zero. Thus, Toda-Yamamoto test could be estimate a VAR with $(k + d_{max})$ order, where d_{max} is the maximal order of integration and k the optimal lagged coefficient jointly. For example, Toda-Yamamoto augmented Granger causality method for two time-series is calculated with the following equations

$$y_t = \alpha_x + \sum_{i=1}^{k+d} \beta_i y_{t-i} + \sum_{j=1}^{k+d} \gamma_j x_{t-j} + u_{xt}, \tag{2.14}$$

$$x_t = \alpha_y + \sum_{i=1}^{k+d} \delta_i x_{t-i} + \sum_{j=1}^{k+d} \epsilon_j y_{t-j} + u_{yt},$$

where u_{xt} and u_{yt} are assumed to be white noise with zero mean and constant variance. If we want to see if x_t affects y_t, we have to test the null hypothesis

$$H_0 : \gamma_1 = \gamma_2 = \ldots = \gamma_j = 0,$$

$$H_a : \text{Some } \gamma_j \text{ is not zero,}$$

If the null hypothesis is rejected then x_t causes y_t. Next alternative null hypothesis test reverses the influential direction

$$H_0 : \epsilon_1 = \epsilon_2 = \ldots = \epsilon_j = 0,$$

$$H_a : \text{Some } \epsilon_j \text{ is not zero,}$$

As this null hypothesis is rejected, we can say that y_t causes x_t. Thus, Granger causality could be determined when time-series are stationary or non-stationary.

Hovewer, it is important to remark the key requirement of any Granger causality test, which is separability, namely that information about a causative factor is independently to that variable and can be taken out by eliminating the variable from the model (as indicated on the right side of equation 2.1). When this requirement is violated Granger causality calculations are no longer valid. Trying to solve this problem, Sugihara et al. (2012) proposed a new causal test, which

named Convergent Cross Mapping (CCM) that is based on the theory of dynamical systems and can be applied to systems where causal variables have synergistic effects.

2.2 Convergent Cross Mapping

CCM is related to a simplex projection which forecast a point in time $t + 1$ for a time-series x, by using points with similar histories from a variable y. This technique is based on Takens' theorem (Takens, 1981), that showed how lagged variables of a single processes can be used as proxy variables to reconstruct an attractor for a dynamic model. Using the last idea, Deyle and Sugihara (2011) expanded the concept of *state space reconstruction (SSR)* to the analysis of the complex, nonlinear systems that appear ubiquitous in the natural, including dynamically coupled processes in the reconstructions of an attractor. In an extension of these ideas, Hsieh et al. (2008) introduced a useful method that avoid data limitation and reduces sampling error by combining similar time-series into a long time-series.

Thus, Sugihara et al. (2012) exposed CCM method based on nonlinear state space reconstruction, that can distinguish causality from correlation. Besides, Sugihara suggest the use of CCM in situations where Granger causality is known to be invalid. Several publications have appeared in recent years documenting implementations of CCM. One of the first examples of applications of CCM is presented in Heskamp et al. (2013), where the authors explored the usability of CCM as a nonlinear analysis technique to assess cerebral regulation and changes in arterial blood pressure. Meanwhile, Luo et al. (2014) use the nonlinear nature of social interactions and adopt CCM to perform causal inference in online social networks on three different social media websites, including Twitter, Sina Weibo (a popular microblog website in

China) and Digg. The latter approach allows authors to infer a causal relationship for any pair of users based on the user activity. Moreover, Tsonis et al. (2015) used CCM to measure causal link between solar activity measured as galactic cosmic rays and year-to-year changes in global temperature. At the same time, Maher and Hernandez (2015) present the first publicly available, open source implementation of CCM (as an official Julia package[1]). Maher and Hernandez (2015) mention some applications as personalized microbiome therapy and metabolic dynamics analysis.

The above mentioned studies show that CCM calculations can successfully uncover the underlying causal structure in each systems. Nonetheless, McCracken and Weigel (2014) show that the relationships between CCM correlations proposed by Sugihara et al. (2012) do not, in general, provide consistent qualification of an intuitive notion of causality in a system. The examples given by the authors cannot lead a conclusion about causality that agree with intuition in the system. In those cases, causality can depend on system parameters. Besides, Mønster et al. (2016) found that CCM could fail to infer causality when the coupling is so strong or weak. And they showed that the presence of noise reduce fidelity in cross-mapping but not convergence.

In economics, testing causality among variables is one of the most important issues. We tested relationship between government expenditure and economic growth in Mexico from 1980 to 2015.

---

[1]`http://pkg.julialang.org/detail/CauseMap.html`

2.3 Causal relationship between government expenditure and economic growth

The relationship between government expenditure and economic growth has been studied by economists and econometricians in very different frameworks. A large number of empirical studies have tried to explain the relation between both variables. In this context, two schools of thought expose the causality of this relationship. On the one hand, Wagner's Law state that a increase in economic growth has consequences (positives or negatives) in government expenditure. On the other hand, Keynesian view conclude that causality takes place from government expenditure to economic growth.

Overall, these studies highlight the need to research models to test Wagner's Law and Keynesian view, in fact these models can be defined by the following relationships:

$$lnGPOP_t = \alpha_0 + \alpha_1 lnYPOP_t + \epsilon_t, \tag{2.15}$$

$$lnYPOP_t = \alpha_0 + \alpha_1 lnGPOP_t + \epsilon_t, \tag{2.16}$$

The specification in Equation 2.15 captures the Wagner's Law and Equation 2.16 captures the Keynesians view. Where ln is the natural logarithm, $GPOP_t$ is real total government expenditure, $YPOP_t$ is the real Gross Domestic Product (GDP) and ϵ_t is the error term at period t.

Many publications have appeared in recent years documenting the dynamic interation

between government expenditure and economic growth. For example, Landau (1983) used cross-country analysis to examine the relatioship between the share of government expenditure in Gross Domestic Product (GDP). The sample available includes over 100 countries including Mexico and show the negative influence of government expenditure on economic growth. In addition, the focus of recent research has been on studies that use time-series data. A key limitation of this research is that a large number of time series are non-stationary. In general, regression models for non-stationary variables give spurious results. To solve this problem, many researchers have applied methods, such as the Granger causality test (Granger, 1969) and cointegrated processes to estimate the long-run effects of economic growth for developed countries. For example Kolluri et al. (2000) and Ghali (1999) used those techniques to evaluate the dynamic between government spending and economic growth for G7 countries and OECD countries, respectively.

The previous studies take into account the direction of causality in developed countries, but little attention has been paid to developing countries. An interesting approach to this issue has been proposed by Samudram et al. (2009), who tested the causal relationship between total expenditure including expenditure of defense, education, development, and agriculture and GDP for Malaysia. They found that with a structural break in 1998, the long-run causality is bi-directional, supporting both Keynesian view and Wagner's Law. In contrast to previous studies with time series, Samudram et al. (2009) added the Auto-Regression Distributed Lag (ARDL) model and the bound test (Pesaran et al., 2001).

In the same spirit, Ono (2014) used the auto-regressive distributed lag test for threshold cointegration, which is developed by Li and Lee (2010). The empirical results indicate that

Wagner's view only holds for annual data of Japan. It is important to note that these previous studies used existing tests to evaluate non-stationary series: Dickey and Fuller (1979), Phillips and Perron (1988), Kwiatkowski et al. (1992) and Kapetanios (2005), cointegration test and error correction: Engle and Granger (1987) and Johansen (1988) and causality: Granger (1969). Several of these tests are included in the category of linear models, besides Granger causality test requires that information about a causative factor is independently unique to that variable.

Chapter 3

Convergent Cross Mapping

Recently there have been an exponential growth in use of social media sites as platforms to build social networks or social relations. This gives a great opportunity to researchers to analyse large-scale data and forecast social trends. As people are influenced by the decision of friends in a social network, the knowledge of who influenced whom has implications in the way how control or promote the spread information or predict social behaviour. If the structure of social network is known, we are able to predict social behaviour, but the causal structure in this system is usually unknown or it is defined as a complex non-linear system. With characteristics above of social network, causal inference in social media could lead a bias estimation of causality. Then CCM will perform causal inference in social media as Luo et al. (2014) developed in their manuscript.

Sugihara et al. (2012) develop a novel method, based on non-linear state space reconstruction of time-serie data. This method called *Convergent Cross Mapping (CCM)* is an alternative to other methods that detect causality between two time-series, basically Granger Causality test. Since Granger causality test is designed to stochastic data, CCM is formulated to measure

correlation in weakly coupled variables belong to deterministic dynamic system. The synthesis of CCM was done according to the procedure of Takens' theorem (Takens, 1981), which ensures that the dynamics of a system can be reconstructed in a model fashion, by using time-delayed embedding to reconstruct its tractor landscape.[1] For example, suppose a two dimensional system $(x - y)$ can be described by its attractor, that is, the trajectory consisting of consecutive equal spaced points. Then, we can use Takens' theorem to reconstruct the attractor from one variable (x) only using time-delayed embedding of the points in the other time-series (y). Sugihara referred to the reconstructed attractor as *shadow manifold*. Thus CCM method is based on the theory of non-linear state space reconstruction (SSR), where SSR techniques have been explored by Deyle and Sugihara (2011). This SSR generalizes Takens' results by improving the reconstruction, when adding multiple dynamical coupled time-series.

For two variables X and Y that are dynamically coupled, a general property of lagged-coordinate embedding is that point $\underline{x(t)}$ on M_x map 1:1 to points $\underline{m(t)}$ on M and local neighbourhoods on M_x map to local neighbourhoods in M, where E is the dimension of a state space, M is the common manifold and the time lag τ is positive. In this way, CCM is an effective way to determine how local neighbourhoods on M_x correspond to local neighbourhoods on M_y. In order to do so, a manifold M_x is constructed by lagged time-series from variable X and used to estimate contemporaneous values of Y. Suppose an attractor

---

[1] In other words,

Theorem 3.0.1. *Let M be a compact, invariant, smooth manifold of dimension m. For pairs (φ, h), where $\varphi : M \to M$ is a smooth observation function (at least C^2) and $h : M \to \mathbb{R}$ a smooth map, the following delay reconstruction map defined by $\Phi[\varphi, h] : M \to \mathbb{R}^{2m+1}$*

$$x \longmapsto (h(x), h(\varphi(x)), \ldots, h()\varphi^{2m}(x)),$$

is an immersion.

manifold M defined by three different variables, its reconstructed attractor or shadow manifolds are defined by $x(t)$, $y(t)$ and $z(t)$ as show Figure 3.1b

(a) Attractor Manifold M **(b)** Shadow Manifolds M_x and M_y
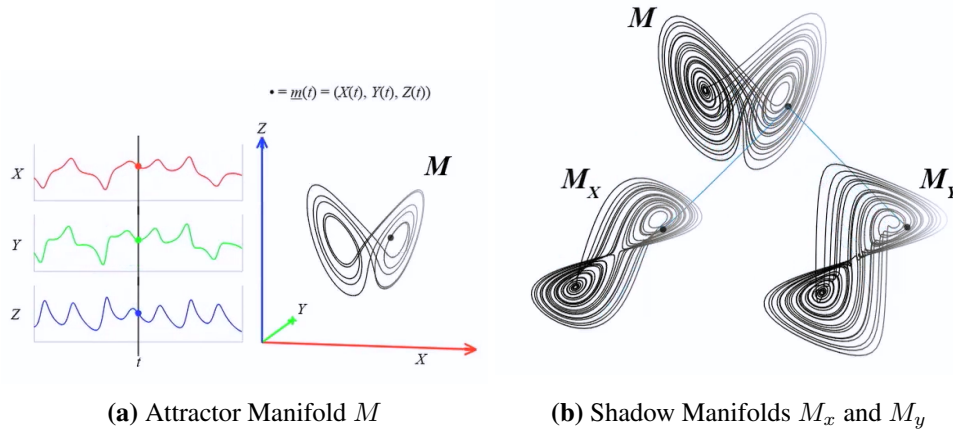
Figure 3.1: (a) Attractor manifold M based on a canonical Lorenz system. A point $m(t)$ on M is defined by $X(t)$, $Y(t)$ and $Z(t)$. **(b)**. Shadow manifolds M_x and M_y. Each point on the shadow manifold M_x is defined by $x(t)$ and $y(t)$ on shadow manifold M_y. Supplementary Materials Sugihara et al. (2012) [2]

3.1 Algorithm of Convergent Cross Mapping

Consider two time series of length L, $\{x\} = \{x_1, x_2, \ldots, x_L\}$, $\{y\} = \{y_1, y_2, \ldots, y_L\}$, and embedding dimension E and a positive time lag τ. Then, x and y can be used to construct shadow manifolds M_x and M_y that are an approximation to the real attractor. The last result was noticed by Takens (1981), considering that shadow manifolds are part of a same dynamical system as a dimension in the state space.

Then, the CCM algorithm may be written in terms of following steps:

---

1. The shadow manifold M_x is reconstructed from $\{x\}$. So, we begin by constructing the lagged-coordinate vectors $\underline{x(t)} = \langle x_t, x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-(E-1)\tau} \rangle$ for $t = 1 + (E-1)\tau$ to $t = L$. Those lagged-coordinate vectors are the base for construction of the shadow manifold M_x.

2. To create a cross-mapped estimate of y_t, denoted by $\widehat{y}_t \mid M_x$. We begin by finding the contemporaneous lagged-coordinate vector on M_x, $\underline{x(t)}$ and locate its $E + 1$ nearest neighbours. For each $\underline{x(t)}$, the nearest neighbour search results in a set of distances that are sorted from closest to farthest by an associated set of time $\{t_1, \ldots, t_{E+1}\}$. Where, the distance $d[\underline{x(t)}, \underline{x(s)}]$ is measured by the Euclidean distance between two vectors.

3. The point in this step is to calculate with a weighted mean the nearest neighbours in M_y. Then, each of the E+1 nearest neighbours are be used to calculate an associated weight. The weight is determined by

$$w_i = \frac{u_i}{\sum_{j=1}^{E+1} u_j},$$

where

$$u_i = exp\{-d[\underline{x(t)}, \underline{x(t_i)}]/d[\underline{x(t)}, \underline{x(t_1)}]\},$$

4. The set of time $\{t_1, \ldots, t_{E+1}\}$ are used to detect neighbours in y. Thus, it is possible to estimate y_t) from a locally weighted mean of the $E + 1$ y_{t_i} values

$$\widehat{y}_t \mid M_x = \sum_{i=1}^{E+1} w_i y_{t_i},$$

Steps 2 to 4 explain how Convergent Cross Mapping reconstructs shadow manifold M_x. Once constructed this manifold, it is necessary to take points that are nearby in M_x and

calculate a weighted mean. The latter mean will ensure a correspondence on M_y and points that are nearby.

5. The CCM correlation is the Pearson correlation coefficient between the original time-series and estimated time-series $\widehat{y}_t \mid M_x$:

$$\rho_{CCM_{YX}} = \rho(y_t, \widehat{y}_t \mid M_x),$$

then we said " y_t **CCM cause** x_t " if correlation $\rho_{CCM_{YX}}$ is not zero. Then, we test the null hypothesis of no correlation against the alternative that there is a non-zero correlation,

$$H_0 : \rho_{CCM_{YX}} = 0$$

$$H_a : \rho_{CCM_{YX}} \neq 0,$$

We calculate a t-statistic for correlation coefficient as,

$$t = \frac{\rho_{CCM_{YX}}}{s_\rho} \text{ where } s_\rho = \sqrt{\frac{1 - \rho^2_{CCM_{YX}}}{N - 2}},$$

where N is the length of the time-series processes. We compute $t_{critical}$ value for Pearson's correlation using a Student's t distribution with $N - 2$ degrees of freedom and a level of significance α, regularly $\alpha = 5\%$. We said $\rho_{CCM_{YX}}$ is not zero, if $\mid t \mid < t_{critical}$, i.e., " Y **CCM cause** X ".

All steps above could be useless if we estimate one point alone because this point is not sufficient to show how well $\widehat{y}_t \mid M_x$ estimates the true value y_t

(a) Shadow Manifolds M_x **(b)** Convergent Cross Mapping
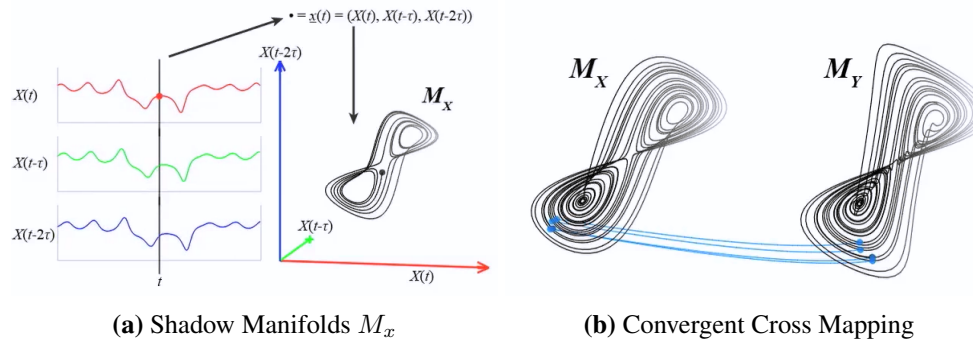
Figure 3.2: **(a)** Shadow manifold M_x constructed using lagged-coordinate embedding of X with $lag = \tau$. **(b)**. Convergent Cross Mapping tests for correspondence between shadow manifolds M_x and M_y. Points that are nearby in M_x will correspond to points that are nearby on M_y. Supplementary Materials Sugihara et al. (2012) [3]

Cross mapping from y_t to x_t is determined analogously. If x_t is affected by y_t, the nearest neighbours of M_x should identify the time indices of corresponding nearest neighbours on M_y. Then, it is necessary obtain a library consisting of L points from M_x is therefore used to provide estimates of L points in the time series for y_t.

---

[3]Source: `http://science.sciencemag.org/content/suppl/2012/09/19/science.1227079.DC1`

Chapter 4

Simple Examples

4.1   Simulations

The usefulness of the CCM method in identifying causative variables into sets of time-series data can be proved by using example systems. Then, the following model examples compare with different sample sizes (T = 50, 200 and 400) the effectiveness of Granger causality tests and CCM to detect causality. The Granger causality tests used in those examples are Granger causality test with vector autoregression model (GC test) and Toda-Yamamoto test (TY test). The latter test was formulated by Toda and Yamamoto (1995) and it can measure Granger causality when time-series variables can be non-cointegrated or cointegrated of an arbitrary order.

The CCM algorithms are implemented in MATLAB, where CCM (Central Matlab) is a file created by Jozef Jakubik and it is available in `http://www.mathworks.com/matlabcentral/`

`fileexchange/52964-convergent-cross-mapping` and CCM (own creation) is an function based on CCM algorithm described in Chapter 3 and is shown in Appendix A. The CCM algorithm depends on the dimension E and the lag time step τ. In the examples we use $E = 3$ and $\tau = 1$. We simulate each systems and test the causal relation between every single processes in the system. Repeating for 1000 times the simulation, we obtain the proportion of simulations when a time-series process causes another variable. On one hand, null hypothesis in Granger causality tests were rejected with a significance level $\alpha = 5\%$. On the other hand, null hypothesis about ρ_{CCM} was rejected with a significance level $\alpha = 5\%$.

Before proceeding to examine causality results, it will be necessary to mention the results obtained by Sugihara et al. (2012). Suppose, a non-linear system for two coupled difference equations that exhibit chaotic behaviour:

$$x_{t+1} = x_t\{r_x - r_x x_t - \beta_{x,y} y_t\},$$
$$y_{t+1} = y_t\{r_y - r_y y_t - \beta_{y,x} x_t\}. \tag{4.1}$$

In Figure 3 of their original text, Sugihara et al. (2012) summarize their results to Equation 4.1 for different values of the coupling constants $\beta_{x,y}$ and $\beta_{y,x}$ with specific values of x_t and y_t. This figure first shows how CCM converges where the effect of x_t on y_t is stronger than in the reverse i.e, $\beta_{y,x} > \beta_{x,y}$, consequently, cross mapping x_t using M_y converges faster than cross mapping y_t using M_x. Secondly, the figure shows that cross mapping of y_t using M_x fails when $\beta_{x,y} = 0$ and it success in cross mapping x_t using M_y. Finally, Sugihara et al. (2012) demonstrated non-convergence of cross mapping y_t using M_x as a function of forcing strength when $\beta_{x,y} = 0$, as a result convergence only occurs as a special case if strong forcing causes the system to collapse dimensionality.

To investigate this further, we will look at what happens with causality for a particular choice of r_x and r_y when the coupling between the two variables is manipulated.

4.1.1   Non-linear independent processes

The first experimental result of Equation 4.1 will be when $\beta_{x,y} = \beta_{y,x} = 0$, and r_x and r_y have different values around but $r_x + r_y = 6.5$. Then, consider the non-linear example of independent processes

$$x_{t+1} = x_t(r_x - r_x x_t),$$

$$y_{t+1} = y_t(r_x - r_x y_t), \tag{4.2}$$

with $r_x + r_y = 6.5$ and starting points x_1 and $y_1 \in [0, 1]$. [1]

Tables 4.1 and 4.2 present the results obtained from simulations for a system where both variables are independent and non-linear with behaviour as a logistic function. On one hand, what is interesting in this system is that CCM methods calculate different percentages of cases where null hypothesis is rejected. This problem may be caused by some ambiguity as Sugihara et al. (2012) illustrated in Figure 3. Nevertheless, causality from x_t to y_t is present in approximately 20% of cases when is used Toda-Yamamoto test. Also CCM (Matlab Central) obtains low percentages of causality.

---

[1]Matlab code file (.m) is available in Appendix B

Table 4.1: Does x_t cause y_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.39	0.22	0.50	0.82
200	0.53	0.19	0.29	0.82
400	0.54	0.15	0.19	0.82

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable X causes Y process. Considering a system with non-linear independent variables (4.2), the most striking result to emerge from the system is that Toda-Yamamoto test rejects null hypothesis in few cases compared with CCM algorithms.

On the other hand, Table 4.2 shows results similar to Table 4.1 CCM algorithms do not reject null hypothesis at least 50% of iterations. Meanwhile, Toda-Yamamoto test rejects null hypothesis in few cases.

Table 4.2: Does y_t cause x_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.40	0.16	0.66	0.80
200	0.35	0.09	0.53	0.79
400	0.38	0.07	0.47	0.76

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable Y causes X process. Considering a system with nonlinear independent variables (4.2), the most striking result to emerge from the system is that Toda-Yamamoto test rejects null hypothesis in few cases compared with CCM algorithms.

Sugihara et al. (2012) demonstrated that cross mapping may fail when $\beta_{x,y}$ is zero and $\beta_{x,y} = 0$ to estimate causality. In Figure 3 from original manuscript, $\rho_{CCM_{YX}} - \rho_{CCM_{XY}}$ has a value between zero and 0.2, if difference is 0.2 then CCM calculated a causality between variables.

4.1.2   Non-linear causal processes

Sugihara et al. (2012) demonstrated the phenomenon of mirage correlation with three different samples from a non-linear logistic difference system as Equation 4.1 with constant coefficients: $r_x = 3.8, r_y = 3.5, \beta_{x,y} = 0.02, \beta_{y,x} = 0.01$ and staring conditions $x_1 = 0.4$ and $y_1 = 0.2$. They illustrate this mirage correlations in Figure 1 of original manuscript and showed that there is no long-term correlations with a length of 1000 in each time-series. In this exercise we want to know if causality changes with this specification of parameters and different starting conditions. Then, consider the non-linear example of a causal processes

$$x_{t+1} = x_t[3.8 - 3.8x_t - 0.02y_t],$$
$$y_{t+1} = y_t[3.5 - 3.5y_t - 0.1x_t], \tag{4.3}$$

with x_1 and $y_1 \sim [0, 0.4]$. This example system is used by Sugihara et al. (2012) to analyse Granger causality tests and how separability affects to tests results. [2]

The results obtained from Equation 4.3 are displayed in Tables 4.3 and 4.4. On the one hand, Granger causality tests identify causality from x_t to y_t almost 90% of iterations, except

---

[2]Matlab code file (.m) is available in Appendix B

Toda-Yamamoto test for time-series with a length $L = 50$ where less than 50% of iterations reject null hypothesis. Meanwhile, CCM reject null hypothesis about $\rho_{CCM} = 0$ almost 70% of samples when time length is $L = 50$, in this case convergence is reflected when the length of a time-series increases.

Table 4.3: Does x_t cause y_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.99	0.47	0.72	0.93
200	1.00	0.97	1.00	1.00
400	1.00	1.00	1.00	1.00

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x_t causes y_t process. Considering a system with non-linear causal variables (4.3), the most striking results to emerge from the system is that Granger causality with Toda-Yamamoto test is present when using small samples in less than 50% of the cases and CCM demonstrates *CCM causality* from x_t to y_t in a higher percentage of samples tested.

On the other hand, Granger causality test does not reject its null hypothesis of y_t causes x_t at least 50% of sample tested, meanwhile Toda-Yamamoto test infers that this causality only 5% of iterations. However, CCM detects causality in an increasing percentage as L grows. While sample length is larger the CCM identifies causality almost every sample.

Table 4.4: Does y_t cause x_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.50	0.05	0.57	0.53
200	0.30	0.04	0.69	0.90
400	0.42	0.05	0.97	1.00

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable Y causes X process. Considering a system with non-linear causal variables (4.3), the most striking result to emerge from the system is that Granger causality is present when using large samples in less than 50% of the cases. The latter result does not agree with intuition about the system created. In this system, separability is not allow then Granger causality calculations are no longer valid. Meanwhile, CCM identifies causality almost every sample.

Sugihara et al. (2012) illustrated this example with samples of 300 points, their results are that *Granger causality using F-test* appears to identify causality from x_t to y_t but it does not detect causality from y_t to x_t. This may due to the parametrization of the system 4.3, which use a $r_y = 3.5$ that is not in the range of chaotic dynamic. In comparison with growth rate $r_x = 3.8$ where the chaos is present. In this case, a chaotic system has a strange attractor, around which the system oscillates forever, never repeating itself or settling into a steady state of behaviour.

4.1.3 Non-linear causal system

Sugihara et al. (2012) showed that CCM can distinguish true interaction from a simple correlation resulted by shared driving variables. In Figure 4 in original manuscript, they illustrated the method with a complex model with five variables (for example, Equation 4.4). They proposed their model such that variables 1, 2, and 3 represent a mutually interacting guild

that externally force variables 4 and 5, whereas 4 and 5 do not influence any other variables. Thus, their results for the complex system indicate that variables 1, 2 and 3 all interact mutually but interact only asymmetrically as external forcing variables with respect to variables 4 and 5, which do not interact directly themselves. We consider the next non-linear example of a causal system with five variables with the same parameters as Sugihara et al. (2012) i.e,

$$x_{1,t+1} = x_{1,t}[4 - 4x_{1,t} - 2x_{2,t} - 0.4x_{3,t}],$$

$$x_{2,t+1} = x_{2,t}[3.1 - 0.31x_{1,t} - 3.1x_{2,t} - 0.93x_{3,t}],$$

$$x_{3,t+1} = x_{3,t}[2.12 + 0.636x_{1,t} + 0.636x_{2,t} - 2.12x_{3,t}],$$

$$x_{4,t+1} = x_{4,t}[3.8 - 0.111x_{1,t} - 0.011x_{2,t} + 0.131x_{3,t} - 3.8x_{4,t}],$$

$$x_{5,t+1} = x_{5,t}[4.1 - 0.082x_{1,t} - 0.111x_{2,t} - 0.125x_{3,t} - 4.1x_{5,t}], \tag{4.4}$$

with starting points $x_i(1) \sim [0, 0.4]$ with $i \in \{1, 2, 3, 4, 5\}$. [3]

Firstly, we obtain causal relation between variable x_1 and other variables. Granger causality tests and CCM method reflect causality from variable 1 to variables 2 and 3 in higher percentage of samples, even length of those samples is short. However, Granger causality is present in low percentage of samples when variable 1 influences variables 4 and 5, for example Toda-Yamamoto test only rejects null hypothesis 10% of iterations when sample size is 400 points and causality is from time-series 1 to 5. Meanwhile CCM causality rejects hypothesis ($\rho = 0$) at least 77% of iterations. This result is consistent with intuition and Sugihara et al. (2012) causal links.

---

[3]Matlab code file (.m) is available in Appendix C

Table 4.5: Does x_1 cause x_j?

Sample size	GC test				TY Test				CCM (Matlab Central)				CCM (own creation)			
	x_2	x_3	x_4	x_5	x_2	x_3	x_4	x_5	x_2	x_3	x_4	x_5	x_2	x_3	x_4	x_5
50	1.00	1.00	0.54	0.56	0.74	0.73	0.11	0.13	1.00	1.00	0.43	0.41	0.95	0.87	0.25	0.25
200	1.00	1.00	0.33	0.27	0.74	0.76	0.13	0.10	1.00	1.00	0.50	0.49	0.99	0.99	0.39	0.53
400	1.00	1.00	0.33	0.26	0.76	0.75	0.16	0.10	1.00	1.00	0.76	0.77	1.00	1.00	0.77	0.93

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x_1 causes x_j processes. Considering a system with non-linear causal variables (4.4), the most striking result to emerge from the system is that Granger causality ($x_1 \Rightarrow x_5$) is present when using large samples in less than 30% of the cases. The latter result does not agree with intuition about the system created but CCM method measures better correlations. In this system, separability is not allow, then Granger causality calculations are no longer valid.

Secondly, we calculate causal relation between variable x_1 and other variables. The results in Table 4.6 reflect the relation from variable 2 to variables 1 and 3. But, as results in Table 4.5, relations from time-series 2 to time-series 4 and 5 are measured poorly by Granger causality tests but CCM algorithms are better measured this relationship. This result is consistent with intuition and Sugihara et al. (2012) causal links.

Table 4.6: Does x_2 cause x_j?

Sample size	GC test				TY Test				CCM (Matlab Central)				CCM (own creation)			
	x_1	x_3	x_4	x_5	x_1	x_3	x_4	x_5	x_1	x_3	x_4	x_5	x_1	x_3	x_4	x_5
50	1.00	1.00	0.53	0.55	0.66	0.66	0.13	0.12	1.00	1.00	0.43	0.38	0.91	0.97	0.24	0.18
200	1.00	1.00	0.30	0.27	0.73	0.72	0.09	0.10	1.00	1.00	0.85	0.44	1.00	1.00	0.84	0.35
400	1.00	1.00	0.40	0.23	0.74	0.76	0.10	0.09	1.00	1.00	1.00	0.61	1.00	1.0	1.00	0.77

Granger Causality and CCM causality are being tested.   The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x_2 causes x_j processes. Considering a system with nonlinear causal variables (4.4), the most striking result to emerge from the system is that Granger causality ($x_2 \Rightarrow x_5$) is present when using large samples in less than 30% of the cases but CCM measures this relation in higher rates. The latter result does not agree with intuition about the system created. In this system, separability is not allow, then Granger causality calculations are no longer valid.

The results in Table 4.7 also reflect a strong causal relation from variable 3 to variables 1 and 2. Meanwhile, the causal relation from variable 3 to variables 4 and 5 has different rates depended on method. This results are similar to variables x_1 and x_2. It is important to note that Granger causality tests only reflect the interaction between variables x_1, x_2 and x_3 and detect causality between those variables and variables x_4 and x_5 in low percentage of samples.

Table 4.7: Does x_3 cause x_j?

Sample size	GC test				TY Test				CCM (Matlab Central)				CCM (own creation)			
	x_1	x_2	x_4	x_5	x_1	x_2	x_4	x_5	x_1	x_2	x_4	x_5	x_1	x_2	x_4	x_5
50	1.00	1.00	0.52	0.56	0.67	0.80	0.13	0.11	1.00	1.00	0.41	0.42	0.97	0.97	0.26	0.28
200	1.00	1.00	0.28	0.24	0.77	0.78	0.09	0.09	1.00	1.00	0.67	0.44	1.00	1.00	0.77	0.32
400	1.00	1.00	0.34	0.24	0.76	0.78	0.07	0.13	1.00	1.00	0.99	0.50	1.00	1.00	1.00	0.44

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x_3 causes x_j processes. Considering a system with nonlinear causal variables, the most striking result to emerge from the system is that Granger causality ($x_3 \Rightarrow x_5$) is present when using large samples in less than 30% of the cases. The latter result does not agree with intuition about the system created. In this system, separability is not allow, then Granger causality calculations are no longer valid.

Sugihara and co-workers demonstrated that Granger causality test misidentifies the causal network. In table 4.8 this misidentification is present in Granger causality test but Toda-Yamamoto test measures the causality better than CCM algorithms. In those latter results, Toda-Yamamoto rejects null hypothesis about causality between variable x_4 and other variables in the lowest rates. Meanwhile, CCM methods reject null hypothesis about correlation contribution in approximately 30% of samples.

Table 4.8: Does x_4 cause x_j?

Sample size	GC test				TY Test				CCM (Matlab Central)				CCM (own creation)			
	x_1	x_2	x_3	x_5	x_1	x_2	x_3	x_5	x_1	x_2	x_3	x_5	x_1	x_2	x_3	x_5
50	1.00	1.00	1.00	0.50	0.11	0.12	0.11	0.06	0.29	0.31	0.32	0.26	0.31	0.28	0.31	0.23
200	1.00	1.00	1.00	0.20	0.09	0.07	0.09	0.05	0.36	0.35	0.36	0.26	0.32	0.32	0.33	0.24
400	1.00	1.00	1.00	0.17	0.08	0.09	0.08	0.05	0.35	0.37	0.38	0.24	0.31	0.32	0.32	0.23

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x_4 causes x_j processes. Considering a system with nonlinear causal variables, the most striking result to emerge from the system is that Granger causality with Granger causality test ($x_4 \Rightarrow x_j$ when $i \in \{1, 2, 3\}$) is present in each sample. The latter result does not agree with intuition about the system created. In this system, separability is not allow, then Granger causality calculations are no longer valid. In this system x_4 causes only x_5 and CCM method identifies this causality in more than 20 % of cases compared to Toda-Yamamoto test, where causality is reflected in 5% of iterations.

As Table 4.9 Toda-Yamamoto test rejects null hypothesis in lowest percentages. Relation between variable x_5 and other variables is calculated better with CCM methods and Toda-Yamamoto test.

Table 4.9: Does x_5 cause x_j?

Sample size	GC test				TY Test				CCM (Matlab Central)				CCM (own creation)			
	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4	x_1	x_2	x_3	x_4
50	1.00	1.00	1.00	0.51	0.11	0.10	0.11	0.09	0.27	0.26	0.26	0.24	0.24	0.24	0.23	0.22
200	1.00	1.00	1.00	0.23	0.09	0.09	0.09	0.05	0.26	0.26	0.24	0.20	0.23	0.23	0.22	0.21
400	1.00	1.00	1.00	0.25	0.10	0.08	0.08	0.05	0.27	0.28	0.26	0.19	0.25	0.21	0.23	0.18

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x_5 causes x_j processes. Considering a system with nonlinear causal variables, the most striking result to emerge from the system is that Granger causality with Granger causality test ($x_5 \Rightarrow x_j$ when $i \in \{1, 2, 3\}$) is present in each sample. The latter result does not agree with intuition about the system created. In this system, separability is not allow, then Granger causality calculations are no longer valid. In this system x_5 causes only x_4 and CCM method identifies this causality in approximately 20 % of cases compared to Toda-Yamamoto test, where causality is reflected in 5% of iterations.

Sugihara et al. (2012) constructed this complex system to prove CCM causality between the subsystem x_1, x_2 and x_3 and variables x_4 and x_5. Then, this subsystem is the forcing subsystem, that interacts unidirectionally with variables x_4 and x_5. But processes x_4 and x_5 do not interact with each other and do not influence x_1, x_2 and x_3

4.1.4 Linear causal processes

This section of results shows a linear system. First, we expect to know how CCM method identifies causality in a linear system. Second, we will study what effect noise on CCM results. Consider the linear example of causal processes

$$y_t = 1.0 + 1.5y_{t-1} - 0.5625y_{t-2} + 0.215x_{t-1} + \epsilon_{yt},$$

$$x_t = 1.0 + 0.8x_{t-1} + \epsilon_{xt}, \tag{4.5}$$

with ϵ_{xt} and $\epsilon_{yt} \sim N(0,1)$. [4]

From the equations which describe the process, we know that the expected conclusion is "x_t causes y_t". The table below illustrates this idea, because for each sample the conclusion is that "x_t Granger causes y_t" and "x_t CCM causes y_t" with high probability. Table 4.10 illustrate this intuitive idea, Granger causality tests and CCM algorithms detect causality in high percentages of samples. This rate increases as sample size grows given the idea of convergence in cross-mapping.

Table 4.10: Does x_t cause y_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.81	0.87	0.62	0.70
200	1.00	0.85	0.78	0.83
400	1.00	0.99	0.88	0.93

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable X causes Y process. Considering a system with linear causal variables (4.5), the most striking result to emerge from the system is that CCM causality is present when using small samples in less than 70% of the cases.

But CCM results in 4.11 does not reliably reflect the intuitive conclusion in this linear example system. In samples larger than 200 observations $\rho_{CCM_{YX}}$ is not zero in more than 70 per cent of simulations. Meanwhile, Toda-Yamamoto test rejects null hypothesis at least 5% of samples.

---

[4]Matlab code file (.m) is available in Appendix D

Table 4.11: Does y_t cause x_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.51	0.81	0.59	0.53
200	0.20	0.07	0.74	0.83
400	0.21	0.05	0.87	0.96

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable Y causes X process. Considering a system with linear causal variables (4.5), the most striking result to emerge from the system is that CCM causality is present when using large samples in more than 80% of the cases. The latter result does not agree with intuition about the system created.

A similar example was reported by McCracken and Weigel (2014) where CCM correlations may not be reliable measure of "driving" for the following dynamical system,

$$x_t = sin(t),$$

$$y_t = Ax_{t-1} + B\eta_t,$$

where $A, B \in \mathbb{R}$ and $\eta_t \sim N(0, 1)$ and length of 2000 observations. Where CCM results depends on A and B parameters. Another conclusion about this kind of example is that the wrong causal effects may be due to a limitation of the algorithm with respect to noise. Mønster et al. (2016) also demonstrated that when the noise level increases the cross-mapped estimation deteriorates. The latter result may be the reason why CCM algorithms do not detect causality from y_t to x_t, as well as, Granger causality tests do.

4.1.5 Structural breaks in time-series

Testing causality when a time-series has structural breaks is of considerable importance in the analysis of time-series. In the section that follows, it will be calculated causality between variables with structural breaks. Firstly, time-series with independent structural breaks, secondly, causal processes with breaks and finally, another causal system with different structural breaks.

Independent processes with independent structural breaks

Consider the following linear example of independent processes with independent structural breaks

$$
x_t = \begin{cases} \mu_{x1} + \phi_{x1}x_{t-1} + u_{xt} & if \quad t \leq t_0; \\ \mu_{x2} + \phi_{x2}x_{t-1} + u_{xt} & if \quad t > t_0, \end{cases}
$$

$$
y_t = \begin{cases} \mu_{y1} + \phi_{y1}y_{t-1} + u_{yt} & if \quad t \leq t_1; \\ \mu_{y2} + \phi_{y2}y_{t-1} + u_{yt} & if \quad t > t_1, \end{cases} \tag{4.6}
$$

with u_{yt} and $u_{xt} \sim N(0,1)$, μ_{x1}, μ_{x2}, μ_{y1} and μ_{y2} are constant, and ϕ_{x1}, ϕ_{x2}, ϕ_{y1} and ϕ_{y2} $\sim U(0,1)$ [5].

Table 4.12 illustrates causality from x_t to x_t. In this example, CCM algorithms detect a causal relation between those variables at least 50% of samples. However, Toda-Yamamoto test rejects its null hypothesis 10% of iterations when sample size is 200 points. The latter result is better than Granger test which detect causality in approximately 30% of samples.

---

[5] Matlab code file (.m) is available in Appendix E

Table 4.12: Does x_t cause y_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.54	0.77	0.52	0.49
200	0.31	0.10	0.52	0.52
400	0.34	0.08	0.53	0.55

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x causes y process. Considering a system with linear independent variables with structural breaks (4.6), the most striking result to emerge from the system is that Granger causality with Toda-Yamamoto test is present in few cases. However, CCM algorithms detect causality at least 50% of cases.

As results above, causality in this system is better measured by Toda-Yamamoto test with sample sizes larger than 50 points.

Table 4.13: Does y_t cause x_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.55	0.77	0.50	0.54
200	0.31	0.10	0.49	0.53
400	0.32	0.07	0.52	0.48

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable y causes x process. Considering a system with linear independent variables (4.6), the most striking result to emerge from the system is that Granger causality with Toda-Yamamoto test is present in few cases. However, CCM algorithms detect causality at least 50% of cases.

The failure of CCM analysis to give expected conclusions about causal interaction with an

independent variables may be due to a limitation of the algorithm with respect to noise.

Causal processes with breaks

Consider the following linear example of independent processes with independent structural breaks

$$
x_t = \begin{cases}
\mu_{x1} + \phi_{x1} x_{t-1} + u_{xt} & if \quad t \leq t_0; \\
\mu_{x2} + \phi_{x2} x_{t-1} + u_{xt} & if \quad t > t_0,
\end{cases}
$$

$$
y_t = \mu_y + \phi_y x_{t-1} + u_{yt}, \tag{4.7}
$$

with u_{yt} and $u_{xt} \sim N(0,1)$, μ_{x1}, μ_{x2} and μ_y are constant, and ϕ_{x1}, ϕ_{x2}, $\phi_y \sim U(0,1)$. [6]

The expected conclusion with System 4.7 is that x_t drives x_t. Calculations in Table 4.14 demonstrate a causal relation from x_t to y_t when Granger causality tests are calculated, for example, Toda-Yamamoto test concludes that x_t causes y_t in more than 90% of samples with length $L = 50$. Nevertheless, CCM algorithms find a causal relation in about 50% of cases.

---

[6]Matlab code file (.m) is available in Appendix E

Table 4.14: Does x_t cause y_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.87	0.91	0.45	0.50
200	0.89	0.77	0.53	0.59
400	0.89	0.75	0.57	0.59

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable x causes y process. Considering a system with linear independent variables (4.7), the most striking result to emerge from the system is that Granger causality with Toda-Yamamoto test is present when using small samples in more than 90% of the cases. Nevertheless, CCM algorithms find a causal relation in about 50% of cases.

Granger causality tests reject the null hypothesis in few cases, this result give the intuitive conclusion that y_t does not cause x_t. But, CCM method rejects null hypothesis about correlation between estimated time-series and real variables.

Table 4.15: Does y_t cause x_t?

Sample size	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
50	0.43	0.76	0.58	0.71
200	0.21	0.07	0.80	0.85
400	0.18	0.05	0.84	0.87

Granger Causality and CCM causality are being tested. The null hypothesis for Granger cause test and Toda-Yamamoto test is rejected with a level of significance $\alpha = 5\%$. The null hypothesis of no correlation in CCM method is rejected with a level of significance $\alpha = 5\%$. The results obtained from a iteration of 1000 repetitions reflect the percent of iterations when variable y causes x process. Considering a system with linear independent variables (4.7), the most striking result to emerge from the system is that CCM causality is present when using large samples in more than 80% of the cases.

As example above, maybe a combinations of coupling strength and noise level give as result that error calculations in CCM algorithms.

Chapter 5

Practical application

In economics, causality could be viewed from the perspective of policy evaluation. Because causal inferences and causal motivations in economics are motivated by policy questions. For example, the analysis of the determinants of long-run economic growth has occupied a central topic in empirical research in economics. A large and growing body of literature has investigated the linkages between economic growth and a variety of economic policy. Much of the available literature on this causal relations deal with the question of causal relation between government expenditure and economic growth. The questions about causal relation between government expenditure and economic growth will be discussed in this chapter. Firstly, we will analyse data from Mexico for both variables. Secondly, we will apply linear methods to know time-series behaviour and finally detect of causality between both variables.

5.1 Data

Gross Domestic Product (GDP hereafter) is considered the most well-known aggregate measure of total economic production for a country and it is an indicator of economic growth. GDP measure the market value of all goods and services produced by an economy during an specific period of time. Hence, GDP includes goods belong to personal consumption, government purchases, private inventories, construction costs and exports. This indicator is normally measured by a national government statistical agency, in Mexico is measured by Instituto Nacional de Estadística, Geografía e Informática (INEGI hereafter).

GDP is measured each quarter, in some cases this indicator is displayed on an annualized basis. Generally, GDP is calculated in real terms, meaning that the data is adjusted for price changes. Data on GDP in Mexico cover the 1980-2015 time period. But, INEGI has divided the calculation into two different periods, the first was constructed on an annualized basis in 1993, the second is calculated on a constant base in 2008. So that, we constructed with all those information available in INEGI website, a long time-series with a total length of 143 observations.

From GDP data, the time-series process has a mean of 9,435,669.51 with a wide variance (SD = 2,465,037.52). Similar to GDP process, government expenditure has a standard deviation equal to 177,949.66. Table 5.1 shows the summary statistics for both time-series variables.

Table 5.1: Descriptive statistics

Basic statistics	GDP	Government Expenditure
Mean	9,435,669.51	1,188,103.48
Median	9,422,001.02	117,910.54
Min.	5,961,524.57	755,222.64
Max.	14,303,118.64	1,571,854.91
S.D.	2,465,037.52	177,949.66

GDP and government expenditure data are obtained from *Instituto Nacional de Estadística y Geografía (INEGI)* $http://www.inegi.org.mx/$, time-series are from 1980 to 2015. Both variables have a trend component and high variance.

The time-series processes present predictable seasonal patterns. As a consequence, this dynamic makes it hard to interpret the trend in both variables. Applying X-12 ARIMA technique to time-series, calculations give us evidence to reject hypothesis about presence of non seasonal processes. Besides, to stabilize the variance of both time-series, we also transform time-series as logarithms. So, we obtain seasonally adjusted processes and the new time-series are shown in Figures 5.1a and 5.1b.

(a) Seasonally adjusted GDP (1980-2015)

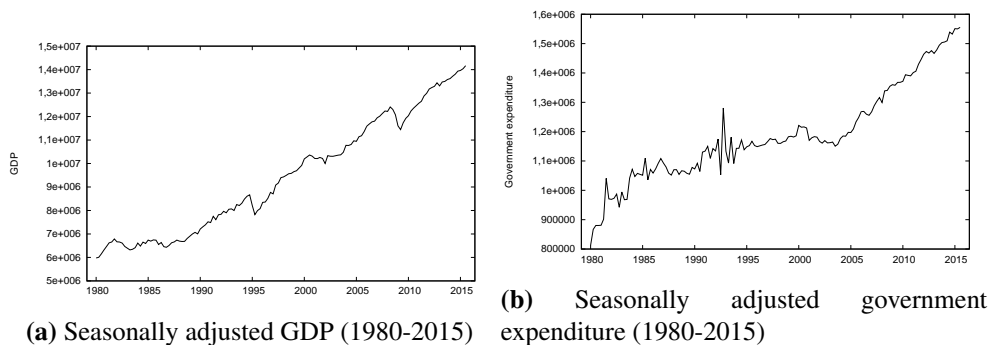**(b)** Seasonally adjusted government expenditure (1980-2015)

Figure 5.1: **(a)** GDP was seasonally adjusted with ARIMA-X12 method in Gretl software. **(b)**. Government expenditure was seasonally adjusted with ARIMA-X12 method in Gretl software.
Source: *Instituto Nacional de Estadística y Geografía (INEGI)* $http://www.inegi.org.mx/$

As demonstrated in the figures above, the seasonally adjusted time-series processes are smoother and shows a trend in time. Trends show the strong growth in both variables. On the one hand, GDP has been growing from the start of 1980 to 2015. On the other hand, government expenditure has grown from 2005 to the present.

Figure 5.1a and 5.1b reveals that after seasonal adjustment, processes still exhibit a nonstationary behaviour. In this case both figures show that the mean and variance are not constant in time. To solve this problem, first, we transform both variables into logarithms, this technique can help to stabilize the variance of a time-series process. And secondly, we compute the first difference, this method help stabilize the mean of a time series by eliminating changes in the level of a time-series.

(a) First difference of GDP (1980-2015)

(b) irst difference of government expenditure (1980-2015)
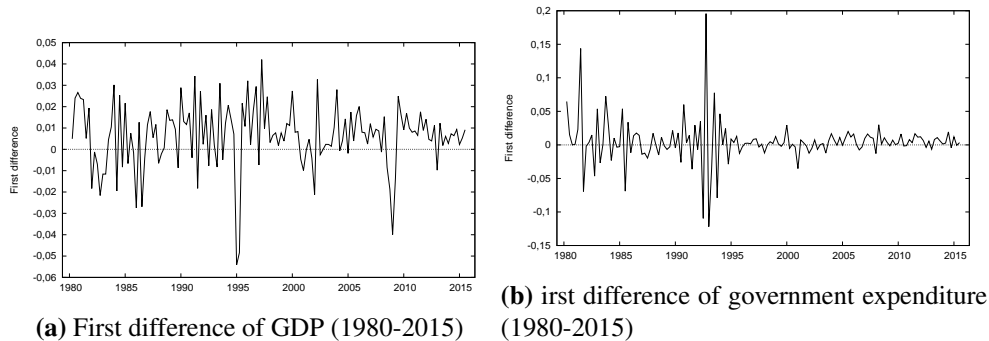
Figure 5.2: **(a)** First difference of GDP time-series obtained with Gretl software. **(b)**. First difference of government expenditure time-series obtained with Gretl software.
Source: *Instituto Nacional de Estadística y Geografía (INEGI)* *http://www.inegi.org.mx/*

5.2 Empirical findings

The results obtained from the preliminary transformation, give us apparently two stationary time-series variables with a standard deviation smaller than data obtained in table 5.1. This

procedure can help to stabilize the mean of time-series by removing changes in the level of a time series, and so eliminating trend.

Table 5.2: Descriptive statistics

Basic statistics	GDP	Government Expenditure
Mean	0.0060	0.0046
Median	0.0076	0.0026
Min.	-0.0541	-0.1219
Max.	0.0420	0.1957
S.D.	0.0156	0.0327

GDP and government expenditure data are obtained from *Instituto Nacional de Estadística y Geografía (INEGI)* $http://www.inegi.org.mx/$, time-series are from 1980 to 2015. First differences are applied for both time-series.

As we showed in Section 2.1.2 unit root tests can be used to determine whether time-series variables are non-stationary. The results obtained from a set of unit root tests tested on logarithmic variables are presented in table 5.3.

Table 5.3: Unit root tests

Variables	Model		ADF	ADF-GLS	KPSS	Philip-Perron
lnGPD						
	Constant	**Level**	0.2511	1.8260	1.2918***	−0.2349
		First difference	−4.1674***	−4.1285***	0.0519	−11.1977***
	Constant and trend	**Level**	−3.0577	−2.3374	0.1099	−2.8140
		First difference	−4.153***	−4.2108***	0.0481	−11.1603***
lnGE						
	Constant	**Level**	0.2571	1.9415	1.2083***	−1.9180
		First difference	−2.7474*	−0.0667	0.1935	−24.3742***
	Constant and trend	**Level**	−1.5208	−1.5667	0.1591**	−5.0276***
		First difference	−2.7993	−1.5861	0.1635**	−24.3288***

*, ** y *** are significant levels in 10%, 5% y 1% respectively to reject null hypothesis

Unit root test for GDP reject hypothesis of stationarity for time-series. It may derive that GDP has an order of integration equal to 1. The order of integration for government expenditure is unknown, all unit root test do not reject null hypothesis.

Source: *Instituto Nacional de Estadística y Geografía (INEGI)* `http://www.inegi.org.mx/`

All these tests use the existence of non stationarity as the null hypothesis, except KPSS where null hypothesis is based on stationarity. It can be seen from the data in Table 5.3 that does not exist enough information to say that both variables are nonstationary. For example, augmented Dickey Fuller and ADF-GLS test test fails to reject the null hypothesis of a unit root test in $lnGDP$. As a result, it is necessary apply other unit root test as Kapetanios test to know if there is a structural break in data.

The nule hypothesis of Kapetanios (2005) is that a time-series process contain a unit root (Theory is explained in section 2.1.3). Kapetanios unit root test has been applied for GDP

and government expenditure both in the first difference. According to the results of Table 5.4, economic growth time-series is stationary in the first difference. Although, in DT model a breaking realized in the data set resulting in 1982.

However, for the government expenditure series the test static at level equal to 5% is greater than the critical value then the basic hypothesis of the time-series that expresses the process contains a unit root can be rejected. Then, time-series data for government expenditure contain more than a break from 1980 to 2015. The break dates are 2004, 1994 and 1983 when government expenditure is modelling with trend dummies, and 2000, 1983 and 1990 when the time-series data is calculating with level dummies and trend dummies.

Table 5.4: Kapetanios test

Variable	Model	1 Break	2 Breaks	3 Breaks	Dates
lnGDP					
	Only DU	-4.5661	-5.0053	-5.8406	$(1996/01, 1984/04, 2008/02)$
	Only DT	-4.4374^{**}	-4.7062	-4.7938	$(1982/03, 2005/04, 2009/02)$
	DU and DT	-4.4231	-5.1083	-5.9847	$(1987/04, 1996/04, 1991/02)$
lnGE					
	Only DU	-4.0105	-4.5279	-4.7639	$(2006/04, 2010/03, 1990/02)$
	Only DT	-4.8093^{**}	-4.8924^{**}	-6.1571^{**}	$(2004/01, 1994/03, 1983/03)$
	DU and DT	-5.2865^{**}	-6.3476^{**}	-6.8665^{**}	$(2000/03, 1983/02, 1990/02)$

** significant level at 5% to reject null hypothesis

Kapetanios test does not reject null hypothesis of unit root in GDP process. For government expenditure data, Kapetanios test reject null hypothesis, then this time-series presents structural breaks in a model whitn trend break dummy variable.

Source: *Instituto Nacional de Estadística y Geografía (INEGI) http://www.inegi.org.mx/*

Table 5.5 shows the summary statistic for a cointegration test using Engle and Granger

(1987) method. This method consists of estimating the cointegration regression by ordinary least square, and applying unit root test for residuals. With these time-series processes, the cointegration test shows that there is no cointegrating vector that makes the variables have the same order of integration. In other words, Engle-Granger method cannot reject null hypothesis of unit root for residuals. This result is obtained by the fact that government expenditure data have structural breaks.

Table 5.5: Cointegration test using Engle-Gragner method

	Without constant	Constant	Constant and trend
ADF test for residuals	-0.6552	-2.2365	-2.6795

Engle-Granger method cannot reject null hypothesis of unit root for residuals. This result is obtained by the fact that government expenditure data have structural breaks.
Source: *Instituto Nacional de Estadística y Geografía (INEGI)* $http://www.inegi.org.mx/$

Johansen et al. (2000) developed a cointegration analysis with a set of variables in the presence of breaks at known points in time. The method is a slight generalization of the likelihood cointegration analysis in vector autoregressive models suggested by Johansen (1988). Table 5.6 presents Johansen test result with structural breaks in government expenditure obtained by Kapetanios test. In table 5.6 rank means order of cointegration, i.e, $rank = 0$ means not cointegration among variables and $rank = 1$ means one cointegration among variables. Then, null hypothesis about no cointegration is rejected using structural breaks, it means that one cointegration vector is possible between those two variables.

Table 5.6: Johansen test trace with breaks

Model	Breaks	Rank	
		0	**1**
Constant			
	2000/03	27.81**	5.13
	(2000/03, 1983/02)	29.82**	5.45
Constant and trend			
	2000/03	25.45**	7.57
	(2000/03, 1983/02)	81.14***	22.31**

*, ** y *** significant level for 10%, 5% y 1%, respectively

One cointegration vector is possible between those two variables in Johansen test with breaks.
Source: *Instituto Nacional de Estadística y Geografía (INEGI)* `http://www.inegi.org.mx/`

As conclusion, government expenditure and economic growth are cointegrated, so both variables have in long-run association-ship. Those two variables are moving together into long-run.

5.3 Granger Causality and CCM Causality

We know that government expenditure and economic growth are non-stationary variables, even government expenditure presents structural breaks in its trend. Engle-Granger cointegration test fails to detect cointegration between those variables, but a modified Johansen test method suggests a cointegration relationship. It is now possible to test causality between both variables. As results in chapter 4, we calculate Granger causality and CCM causality to predict causality direction between economic growth and government expenditure.

First, we test these variables in level null hypothesis in Granger causality test is rejected, i.e. economic growth **Granger causes** government expenditure. The null hypothesis in CCM causality is rejected in both sides of causal relationship and correlation in both sides are almost one. Based on results calculated in a system containing variables with structural changes CCM methods cannot predict causal direction. In this example, CCM method gives ambiguous relations.

Table 5.7: Does GDP cause Government expenditure?

Causal relationship	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
GDP ⇒ Government expenditure	4.42**	0.76	0.97**	0.98**
Government expenditure ⇒ GDP	1.58	0.71	0.97**	0.94**

Null hypothesis in Granger causality test are tested with a significant level $\alpha = 5\%$. Test static at level equal to 5% and $t_{critical}$ value for Pearson's correlation with $N - 2$ degrees of freedom and $\alpha = 5\%$ of significance. With Granger causality test, we know that GDP drives government expenditure.
Source: *Instituto Nacional de Estadística y Geografía (INEGI)* `http://www.inegi.org.mx/`

Secondly, we test these first difference variables, null hypothesis in Granger causality tests is rejected, i.e. economic growth **Granger causes** government expenditure. The null hypothesis in CCM causality cannot be rejected in both sides of causal relationship. CCM method cannot detect causal direction because variables with first difference lose prediction to other time-series.

Table 5.8: Does GDP cause Government expenditure?

Causal relationship	GC test	TY Test	CCM (Matlab Central)	CCM (own creation)
GDP $\Rightarrow$ Government expenditure	3.09**	0.92**	-0.05	-0.02
Government expenditure $\Rightarrow$ GDP	1.86	0.59	0.15	-0.08

Null hypothesis in Granger causality test are tested with a significant level $\alpha = 5\%$. Test static at level equal to 5% and $t_{critical}$ value for Pearson's correlation with $N - 2$ degrees of freedom and $\alpha = 5\%$ of significance. With Granger causality tests, we know that GDP drives government expenditure.

Source: *Instituto Nacional de Estadística y Geografía (INEGI) http://www.inegi.org.mx/*

The causal relation could be modelled as,

$$lnGPOP_t = \alpha_0 + \alpha_1 lnYPOP_t + \epsilon_t, \tag{5.1}$$

where ln is the natural log, $GPOP_t$ is real total government expenditure, $YPOP_t$ is real Gross Domestic Product and ϵ_t is the error term.

Equation 5.1 captures Wagner's law. Comín et al. (2009) proved that economic growth causes government expenditure in countries as Argentina, Brazil, Mexico and Spain in twentieth century period. Samudram et al. (2009) showed a long-run relationship runs from GDP to the expenditure in the case of a developing country, Malaysia. Tables 5.7 and 5.8 showed that growth expenditure causes government expenditure for Mexico from 1980 to 2015.

Chapter 6

Conclusions

In science, methods for inference of causality between two or more variables are of great interest. Granger causality test (Granger, 1969) is used to detect causality in easily separable linear systems, Sugihara et al. (2012) intriduced an an alternative to detect causality between two time series when separability is violated. This novel method is named CCM and provides a framework that uses predictability as opposed to correlation to identify causation between time-series. CCM has already been used in a wide range of different fields for different kinds of data (Heskamp et al. (2013); Luo et al. (2014); Maher and Hernandez (2015); Tsonis et al. (2015)) and it has been noted (McCracken and Weigel (2014) and Mønster et al. (2016)) that CCM results are not always consistent with theoretical intuitions and parametrization.

Given the increasing interest and alternative to measure causality by CCM, it is important to understand its strengths and limitations. The present study was designed to present an analysis of simple model systems. It was noted by Sugihara et al. (2012) that CCM does not correctly predict the direction of causality in certain coupling of variables, as we have also seen in Section

4. For example, our analysis shows that CCM method seems sensitive to noise terms and to particular dynamics (i.e., non-linear independent systems where a chaotic system appears for growth rate parameter beyond 3.5. This chaotic system has a strange attractor, around which the system oscillates forever, never repeating itself or settling into a steady state of behaviour. It never hits the same point twice and its structure has a fractal form, meaning the same patterns exist at every scale.) where this method fails to predict the right direction of causality. However, Toda-Yamamoto test betters indicates causal relationships that agree with intuition when noise terms are added. An advantage of using Toda-Yamamoto test is that it can be applied regardless of whether a time-series has order of integration zero or greater than zero, non-cointegrated or cointegrated of an arbitrary order.

The application of CCM to real-world data for purposes of causal inference, in this case government expenditure and economic growth for Mexico (1980-2015). Granger causality tests and CCM methods were used to capture the dynamic between those variables. We also tested for structural breaks in the data and found that government expenditure presents breaks in 1994 maybe as a result of *The Tequila Crisis*. Without structural breaks, the null hypothesis in Engel-Granger test is not rejected. With structural breaks, Johansen test found a cointegrated vector. Then, economic growth and government expenditure are long-run forcing. On one hand, Granger causality test and Toda-Yamamto test indicated directional causality from GDP to government expenditure. These result support Wagner's law. On the other hand, CCM algorithms could not indicate a causal relation given limitations in the method as shown in variables with no-chaotic behaviour, structural breaks and noise term.

Appendices

A CCM algorithm

```matlab
function [Mx, Map, corre] =CCM(X,Y,tau,E)
%% The function CCM computes the method Convergent Cross Mapping Sugihara
% et al. (2012) which is complementary to Granger Causality test.
% Since Granger causality test is designed to stochastic data, CCM is
% formulated to measure correlation in coupled variables belong to
% deterministic dynamic system.
%
% User-Specified Inputs:
%   X: A columm vector of data
%   Y: A column vector of data
%   tau: a postive time lag
%   E: embedding dimension
%
% User-requested Output:
%   Mx: The shadow manifold reconstructed from X
%   Map: a cross-mapped estimate of Y
%   corre: Pearson correlation coefficient between the original timeseries
%   and estimated time-series of Y
%
% Referece:
%   Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M.,
%   and Munch, S. (2012). Detecting causality in complex ecosystems.
%   science, 338(6106):496?500.
%
% Authors:
% Daniel Ventosa-Santaularia and Rubi Tonantzin Gutierrez Villanueva
% CIDE, Spring 2016
L=length(X);
dist1=L-((E-1)*tau);
dist2= E;
%Create a matriz Mx called "Shadow manifold"
Mx=zeros(dist1,dist2);
Map=zeros(dist1,1);
```

```matlab
34    Mx_index=zeros(dist1,dist2);
35    %Reconstructed mainfold
36    for t=1+(E-1)*tau:L
37        for tt=1:E
38            Mx(t-(1+(E-1)*tau)+1,tt)=X(t-(tt-1)*tau);
39            Mx_index(t-(1+(E-1)*tau)+1,tt)=t-(tt-1)*tau;
40        end
41    end
42    %Genereate a cross-mapped estimate of Y(t)
43    [m n] = size(Mx);
44    Es = zeros(dist1,E+1);
45    Dist1 = Es;
46    for tt = 1:m
47        Aux1 = zeros(1,m);
48        % Finding the contemporaneous lagged-coordinate vector on Mx
49        for pp = 1:m
50            Aux1(pp) = norm(Mx(tt,:)-Mx(pp,:));
51            if Aux1(pp)==0
52                Aux1(pp)=inf(1);
53            end
54        end
55        % locating its E + 1 nearest neighbours
56        [B,I] = sort(Aux1);
57        for rr=1:(E+1)
58            Dist1(tt,rr) = B(rr);
59            Es(tt,rr) = I(rr);
60        end
61        % Detecting neighbours in Y
62        ui = zeros(1,E+1);
63        for ii=1:(E+1) % associated weight
64            ui(ii)=exp(-Dist1(tt,ii)/Dist1(tt,1));
65        end
66        Uis = sum(ui);
67        yt=0;
68        for yy=1:(E+1)
69            yt=(ui(yy)/Uis)*Y(((E-1)*tau)+Es(tt,yy))+yt;
70        end
71        Map(tt)=yt;
72
73    end
74    % CCM correlation
75    corre = corr(Y(1+(E-1)*tau:L), Map);
76
77    end
```

B Generating nonlinear causal processes

```matlab
function [X Y] = generador(rx,ry,bxy,byx,y1,x1,L)
%% The function generador computes two coupled difference equations that
% exhibit chaotic behavior:
% X(t+1) = X(t)[rx-rxX(t)-bxyY(t)]...1
% Y(t+1) = Y(t)[ry-ryY(t)-byxX(t)]...2
%
% Inputs:
%    rx: constant and parameter for variable X in equation 1
%    ry: constant and parameter for variable X in equation 2
%    bxy: correlation between variable X and Y in equation 1
%    byx: correlation between variable Y and X in equation 2
%    x1: first value for X
%    y1: first value for Y
%    L: length of time-series processes
%
% Output:
%    X: time-series proceses X
%    Y: time-series proceses Y
%
% Author: Rubi Tonantzin Gutierrez Villanueva
% CIDE, Spring 2016
%
X=zeros(1,L);
Y=zeros(1,L);

X(1)=x1;
Y(1)=y1;

for i=2:L

    X(i) = X(i-1)*(rx-rx*X(i-1)-bxy*Y(i-1));
    Y(i) = Y(i-1)*(ry-ry*Y(i-1)-byx*X(i-1));
end

end
```

C Generating nonlinear causal system

```matlab
function [x1,x2,x3,x4,x5] = generador_sistema(val1,val2,val3,val4,val5,T)
%% The function generador_sistema computes a coupled system that
% exhibit chaotic behavior:
% X1(t + 1) = X1(t)[c1 + r1X1(t) + b12X2(t) + b14X3(t)]...1
% X2(t + 1) = X2(t)[c2 + b21X1(t) + r2X2(t) + b23X3(t)]...2
% X3(t + 1) = X3(t)[c3 + b31X1(t) + b32X2(t) + r3X3(t)]...3
% X4(t + 1) = X4(t)[c4 + b41X1(t) + b42X2(t) + b43X3(t) + r4X4(t)]...4
% X5(t + 1) = X5(t)[c5 + b51X1(t) + b52X2(t) + b53X3(t) + r5X5(t)]...5
%
% Inputs:
%  val1: constant, parameters and first value for variable X1 in equation 1
%  val2: constant, parameters and first value for variable X2 in equation 2
%  val3: constant, parameters and first value for variable X3 in equation 3
%  val4: constant, parameters and first value for variable X4 in equation 4
%  val5: constant, parameters and first value for variable X5 in equation 5
%   T: length of time−series processes
%
% Output:
%   x1: time−series proceses X1
%   x2: time−series proceses X2
%   x3: time−series proceses X3
%   x4: time−series proceses X4
%   x5: time−series proceses X5
%
% Author: Rubi Tonantzin Gutierrez Villanueva
% CIDE, Spring 2016


x1 = zeros(T,1); x2 = zeros(T,1); x3 = zeros(T,1); x4 = zeros(T,1);
x5 = zeros(T,1);

x1(1) = val1(1); x2(1) = val2(1); x3(1) = val3(1); x4(1) = val4(1);
x5(1) = val5(1);



for i=2:T

    x1(i)= x1(i−1)*(val1(2) + val1(3)*x1(i−1) + val1(4)*x2(i−1) + ...
        val1(5)*x3(i−1));
    x2(i)= x2(i−1)*(val2(2) + val2(3)*x1(i−1) + val2(4)*x2(i−1) + ...
        val2(5)*x3(i−1));
    x3(i)= x3(i−1)*(val3(2) + val3(3)*x1(i−1) + val3(4)*x2(i−1) + ...
        val3(5)*x3(i−1));
    x4(i)= x4(i−1)*(val4(2) + val4(3)*x1(i−1) + val4(4)*x2(i−1) +...
        val4(5)*x3(i−1) + val4(6)*x4(i−1));
```

```
47        x5(i)= x5(i-1)*(val5(2)  +  val5(3)*x1(i-1)  +  val5(4)*x2(i-1)  +...
48            val5(5)*x3(i-1)  +  val5(6)*x5(i-1));
49
50   end
51
52   end
```

D   Generating linear causal processes

```matlab
function [X Y] = generador_causal(cx, rx, cy, ry1, ry2, byx, L)
%% The function generador_causal computes two linear causal processes
% x(t) = cx + rxX(t-1) + err(t)...1
% y(t) = cy + ry1Y(t-1) + ry2Y(t-2) + byxX(t-1) + err(t)...2
%
% Inputs:
%    cx: constant for variable y in equation 1
%    rx: parameter for variable X in time t-1 as it is shown in equation 1
%    cy: constant for variable y in equation 2
%    ry1: parameter for variable Y in time t-1 as it is shown in equation 2
%    ry2: parameter for variable Y in time t-2 as it is shown in equation 2
%    byx: correlation between variable Y and X in equation 2
%    L: length of time-series processes
%
% Output:
%    X: time-series proceses X
%    Y: time-series proceses
%
% Author: Rubi Tonantzin Gutierrez Villanueva
% CIDE, Spring 2016
%
eps1 = randn(L,1);
eps2 = randn(L,1);
X = zeros(L,1);
Y = zeros(L,1);

X(1)=eps1(1); Y(1)=eps2(1); Y(2)= eps2(2);

for i = 2:L
    X(i)=cx+rx*X(i-1)+eps1(i);
end

for j=3:L
    Y(j)=cy+ry1*Y(j-1)+ry2*Y(j-2)+byx*X(j-1)+eps2(j);
end

end
```

E   Generating structural breaks in time-series

```matlab
function [y, x] = quiebres(T,n,tipo, phi, theta, mu1, mu2,mu3,mu4)
%% The function quiebres computes two linear causal processes with
% structural breaks
%
% Inputs:
%    T: length of time-series processes
%    n: number of initial conditions
%    tipo: 1 -> Independent time-series
%          2 -> Dependent time-series
%          3 -> Dependent time-series in trend
%    phi: parameters for variable Y
%    theta: parameters for variable X
%    mu1: constant for variable X
%    mu2: constant for variable X
%    mu3: constant for variable Y
%    mu4: constant for variable Y
%
%
% Output:
%    y: time-series proceses y
%    x: time-series proceses x
%
% Author: Rubi Tonantzin Gutierrez Villanueva
% CIDE, Spring 2016
%

uy = randn(T+n);
ux = randn(T+n);


y = zeros(T+n,1);
x = zeros(T+n,1);

a= 10+n;
b = (T+n)-10;

y(1) = uy(1);
x(1) = ux(1);

cambio1 = a +(b-a)*rand(1);
cambio2 = a +(b-a)*rand(1);


%Independent time-series x
for i = 2:T+n
    if i <= cambio2
```

```
47            x(i) = mu1 + theta(2)*x(i-1) + ux(i);
48        else
49            x(i) = mu2 + theta(1)*x(i-1) + ux(i);
50        end
51   end
52
53
54   %Independent time-series y
55   if tipo == 1
56        for i = 2:T+n
57            if i <= cambio1
58                y(i) = mu3 + phi(2)*y(i-1) + uy(i);
59            else
60                y(i) = mu4 + phi(1)*y(i-1) + uy(i);
61            end
62        end
63
64   %Dependent time-series y
65   elseif tipo == 2
66
67        for i = 2:T+n
68            y(i) = phi(1) + phi(2)*x(i-1) + uy(i);
69        end
70
71   %Dependent time-series y
72   elseif tipo == 3
73
74        for i = 2:T+n
75            if i <= cambio2
76                y(i) = phi(1) +  phi(2)*y(i-1) + phi(3)*x(i-1) + uy(i);
77            else
78                y(i) = phi(1) + phi(2)*y(i-1) + uy(i);
79            end
80        end
81
82
83   end
```

Bibliography

Agueda, C. P. (2011). Causality in sciencie. *Pensamiento Matemático*, page 12.

Comín, F. C., Fuentes, D. D., and López, J. R. (2009). La relación entre el crecimiento económico y el gasto público en argentina, brasil, españa y méxico durante el siglo xx. In *XVI Encuentro de Economía Pública: 5 y 6 de febrero de 2009: Palacio de Congresos de Granada*, page 125.

Deyle, E. R. and Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. *PLoS One*, 6(3):e18295.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74(366a):427–431.

Engle, R. F. and Granger, C. W. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276.

Ghali, K. H. (1999). Government size and economic growth: evidence from a multivariate cointegration analysis. *Applied Economics*, 31(8):975–987.

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

Granger, C. W. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics*, 2(2):111–120.

Heskamp, L., Meel-van den Abeelen, A., Katsogridakis, E., Panerai, R., Simpson, D., Lagro, J., and Claassen, J. (2013). Convergent cross mapping: a promising technique for future cerebral autoregulation estimation. In *Cerebrovascular Diseases*, volume 35, pages 15–16.

Hsieh, C.-h., Anderson, C., and Sugihara, G. (2008). Extending nonlinear analysis to short ecological time series. *The American Naturalist*, 171(1):71–80.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of economic dynamics and control*, 12(2):231–254.

Johansen, S., Mosconi, R., and Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *The Econometrics Journal*, 3(2):216–249.

Kapetanios, G. (2005). Unit-root testing against the alternative hypothesis of up to m structural breaks. *Journal of Time Series Analysis*, 26(1):123–133.

Kolluri, B. R., Panik, M. J., and Wahab, M. S. (2000). Government expenditure and economic growth: evidence from g7 countries. *Applied Economics*, 32(8):1059–1068.

Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of econometrics*, 54(1):159–178.

Landau, D. (1983). Government expenditure and economic growth: a cross-country study. *Southern Economic Journal*, pages 783–792.

Li, J. and Lee, J. (2010). Adl tests for threshold cointegration. *Journal of Time Series Analysis*, 31(4):241–254.

Luo, C., Zheng, X., and Zeng, D. D. (2014). Causal inference in social media using convergent cross mapping. In *JISIC*, pages 260–263.

Maher, M. C. and Hernandez, R. D. (2015). Causemap: fast inference of causality from complex time series. *PeerJ*, 3:e824.

McCracken, J. M. and Weigel, R. S. (2014). Convergent cross-mapping and pairwise asymmetric inference. *Physical Review E*, 90(6):062903.

Mønster, D., Fusaroli, R., Tylén, K., Roepstorff, A., and Sherson, J. F. (2016). Inferring causality from noisy time series data. *arXiv preprint arXiv:1603.01155*.

Morck, R. and Yeung, B. (2011). Economics, history, and causation. *Business History Review*, 85(01):39–63.

Ono, H. (2014). The government expenditure–economic growth relation in japan: an analysis by using the adl test for threshold cointegration. *Applied Economics*, 46(28):3523–3531.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica: Journal of the Econometric Society*, pages 1361–1401.

Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, 16(3):289–326.

Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2):335–346.

Samudram, M., Nair, M., and Vaithilingam, S. (2009). Keynes and wagner on government expenditures and economic development: the case of a developing economy. *Empirical Economics*, 36(3):697–712.

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., and Munch, S. (2012). Detecting causality in complex ecosystems. *science*, 338(6106):496–500.

Takens, F. (1981). *Detecting strange attractors in turbulence*. Springer.

Toda, H. Y. and Yamamoto, T. (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics*, 66(1–2):225–250.

Tsonis, A. A., Deyle, E. R., May, R. M., Sugihara, G., Swanson, K., Verbeten, J. D., and Wang, G. (2015). Dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. *Proceedings of the National Academy of Sciences*, 112(11):3253–3256.