

# On the empirical relationship between tourism and economic growth

Juan Gabriel Brida<sup>a,\*</sup>, David Matesanz Gómez<sup>b</sup>, Verónica Segarra<sup>a</sup>

<sup>a</sup> GIDE, Faculty of Economics, University of the Republic, Full postal address: Gonzalo Ramirez 1926, 11200, Montevideo, Uruguay

<sup>b</sup> Department of Applied Economics, University of Oviedo, Full postal address: Avda. del Cristo s/n 33006, Oviedo, Spain

## ARTICLE INFO

### Keywords:

Regime dynamics  
Cluster analysis  
Economic growth  
Tourism development

## ABSTRACT

This paper studies the dynamics of economic growth and tourism evolution for 80 countries during the period 1995–2016. The variables representing economic and tourism growth are growth rates of per capita GDP and international tourist arrivals per inhabitant respectively. Using the concept of economic regime, the paper introduces a notion of distance between the dynamical paths of different countries. Then, a Minimal Spanning Tree and a Hierarchical Tree are constructed to detect groups of countries sharing similar performance. The two main clusters we find can be interpreted as two groups of countries with high and low performance in the tourism sector and are coherent with the business cycle. The evolution of such clusters shows three main *stylized* facts: certain countries move across clusters; the low performance cluster tends to span, while the high performance one tends to be (more) compact; the distance between the two groups increases in time.

## 1. Introduction

In this paper, the relationship between tourism and economic growth is explored. To this end, the study introduces an alternative non parametric methodology to the very well-known econometric tools commonly used in the empirical literature investigating tourism and economic growth. This methodology is used to study the dynamics of economic growth and tourism performance for 80 countries during the period 1995–2016. The fact that the methodology do not need to assume a particular model to study the relationship between the variables under study allows to understand the dynamic behaviour of the countries in the sample and to compare their performance. With this usage, the so-called tourism-led-growth-hypothesis (TLGH) can be seen in a different light and could call for a complete reformulation. In particular, our contribution to the TLGH literature includes the possibility of testing if different countries could admit a similar model representing the dynamics in tourism and economic growth. In addition, this tool can be used when working with panel data to test the homogeneity of individuals in the panel. This central hypothesis of homogeneity of individuals is addressed by using clustering techniques that help to find homogeneous groups of countries with similar dynamics in tourism and economic growth. Then panel data can be correctly used for each cluster in the sample.

The tourism sector is recognized to positively contribute to the economic growth process of a country through different channels,

including the fact of course that it is a currency earner sector; that stimulates physical and human capital accumulation, and pushes (and uses) technology and innovation. At the same time, tourism promotes directly and indirectly other economic industries such as transportation, hospitality or retailing (see Mayer & Vogt, 2016). In particular, international tourism is a source of foreign currency that facilitates the acquisition of capital goods and technologies, which can be used in other production processes. In addition, it plays a significant role stimulating investments in new infrastructure and promoting competition, creating employment and corresponding household income. Last but not least, it must be noted that tourism is a significant sector of diffusion of technical knowledge, and potentially it can stimulate research and development.

Many governments are paying a greater attention to support and promote tourism as a potential source of growth and employment and as a sector which adds value to cultural, natural and other capital with no market price. On the other hand, on the academic sector there is a growing interest in the relationship between tourism and economic growth, particularly from the empirical perspective. However, the theoretical justification of such relationship is still not explicitly defined. The seminal study by Balaguer and Cantavella-Jordà (2002) is recognized to be the seminal paper that formalizes the so-called tourism-led-growth hypothesis (TLGH), therefore offering a theoretical and empirical connection between tourism and economic growth.

Theoretically, the TLGH was directly derived from the export-led growth hypothesis that postulates that economic growth can be

\* Corresponding author.

E-mail addresses: [gbrida@ccee.edu.uy](mailto:gbrida@ccee.edu.uy) (J.G. Brida), [matesanzdavid@uniovi.es](mailto:matesanzdavid@uniovi.es) (D. Matesanz Gómez), [vsegarra@ccee.edu.uy](mailto:vsegarra@ccee.edu.uy) (V. Segarra).

<https://doi.org/10.1016/j.tourman.2020.104131>

Received 16 January 2020; Received in revised form 9 April 2020; Accepted 24 April 2020

Available online 11 May 2020

0261-5177/© 2020 Elsevier Ltd. All rights reserved.

generated and statistically caused by expanding exports. Since that seminal study, hundreds of papers have been published including different case studies, methodologies and databases. There already exist several literature review and meta-analysis studies analysing the TLGH and the relationship between tourism and economic growth (for example, Nunkoo, Seetanah, Jaffur, Moraghen, & Sannassee, 2019; Fonseca & Sánchez Rivero, 2019a; Fonseca & Sánchez Rivero, 2019b; Comerio & Strozzi, 2019; Chingarande & Saayman, 2018; Li, Jin, & Shi, 2018; Brida, Cortes-Jimenez, & Pulina, 2016; Brida, Lanzilotta, Pizzolon, 2016; Seetanah, Nunkoo, Sannassee, Georges, & Jaffur, 2017; Castro-Nuño, Molina-Toucedo, & Pablo-Romero, 2013; Castro-Nuño, Molina-Toucedo, & Pablo-Romero, 2013; Adamou & Clerides, 2010). The greater part of the more than two hundred published case studies analyzed in the previous surveys provide support or the TLGH and report positive and statistically significant estimations. These reviews also suggest that estimations are sensitive to a number of specific variables such as country data, specification, estimation characteristics, and time span. This implication opens a window to review the models used in the analysis and to empirically study the relationship between tourism and economic growth, but without having a particular model in mind and trying to understand what data can tell. The most recent works, in terms of methodological and modelling innovation introduce non-linear analysis and non-parametric tests of causality (see Zhang & Cheng, 2019; Eyuboglu & Eyuboglu, 2019; Gül & Özer, 2018; Karimi, 2018; Bella, 2018; Chiang, Sung, & Lei, 2017; Brida, Lanzilotta, Pereyra, & Pizzolon, 2015; Brida et al., 2016a; Brida et al., 2016b; Kumar & Stauvermann, 2016; Wu, Liu, Hsiao, & Huang, 2016; and references therein). These are recent approaches to the TLGH revealing interesting results that partially contradicts the general conclusions obtained in more traditional linear models and standard causality tests. Finally, some recent papers also focus on global or panel data studies (see De Vita & Kyaw, 2017; Chiu & Yeh, 2017; Tang & Tan, 2015; Fahimi, Saint Akadiri, Seraj, & Akadiri, 2018; Tang & Tan, 2018; Škrinjarčić, 2019; Sokhanvar, Çiftçioğlu, & Javid, 2018; Muslija, Satrovic, & Erbaş, 2017; and Risso, 2018; and references therein). In particular, the present study also contributes to this empirical strand of research. Panel data are multi-dimensional statistics regarding measurements over time that comprise observations of multiple phenomena obtained over multiple time periods for the same entities. But rarely the homogeneity of individuals is tested. In this paper, the central hypothesis of homogeneity of individuals is addressed by using clustering techniques that help to find homogeneous groups of countries with similar dynamics in tourism and economic growth.

The most recent papers use panel data methodology with the disadvantage of considering a single model for all agents in their dataset. In studies that use panel data, homogeneity in the behaviour of the different countries of the panel is not analyzed, checking whether the parameters of the model are common to all individuals. Additionally, models are usually linear except for a few of studies. In general, and particularly from the most recent non-linear and non-parametric exercises, the TLGH literature yield evidence in favour of the view that a unique interpretive model is likely to be inadequate to describe growth and tourism development experiences. However, if we accept diversity across countries and regions, we have to re-conceptualize the dynamic representation of tourism and economic growth. The aim of this paper is to analyse the role of tourism development for a set of countries by analysing the dynamic relationship between tourism performance and economic growth. In particular, we compare their qualitative dynamical behaviour. We aim to contribute in this direction, comparing the dynamic behaviors of the different countries without having in mind any model. In this way, it is possible to identify groups of countries that have a similar dynamic, for which you could search for models of the same type. In this sense, we are directly interested in the heterogeneity presented in the connections between tourism and economic growth and the consequences it implies for the empirical analysis and policy implications. The results of the study show that there are different groups

of countries presenting a similar dynamical behaviour in tourism and growth. These groups are characterized by the level of tourism specialization.

These previous points make explicit the idea that a country's dynamics in tourism and economic growth can be better characterized by employing a new description based on the notion of economic regime (Brida, 2008; Cristelli, Tacchella, & Pietronero, 2015). The proposed notion and the involved methodology advances over the existing literature by imposing an order on the regimes and adopting a more adequate analytical technique, called symbolic time series analysis (Risso, 2017).

Many papers on economic growth and tourism consider the number of arrivals of international tourists or their spending as a measure to quantify the importance of tourism (Rosselló-Nadal & HE, 2019). These studies find important differences in their results depending on the measure of the variable tourism they have used. When we consider the absolute number of international tourists received, part of the phenomenon we want to capture is lost, since the "size" of the economy is not taken into account. For example, in 2016 Brazil received 6.547.000 tourists, and Uruguay received just under half: 3.037.000. Thus, in "absolute" terms, Uruguay receives a smaller number of tourists than Brazil. But what weight does this number have in relation to the size of the economy? In terms of population, in 2016 Brazil had 207.652.865 inhabitants, that is, they received 3 tourists per 100 inhabitants, and the population in Uruguay was 3.444.006 inhabitants, thus receiving almost one tourist per habitant (88 tourists per 100 inhabitants). For this reason, we normalize the variable tourism in terms of the number of inhabitants of each country, in order to collect the "weight" of the number of tourists received, with respect to the size of each country. In this sense, we follow a few number of previous empirical studies employing per capita tourist arrivals, including Chiu and Yeh (2017), among others.

The paper is organized as follows. In section 2 we describe the data and the methodology. Section 3 introduces the concepts of both the minimum spanning tree and the hierarchical tree and presents the results of the empirical application of the methodology. In section 4 we introduce time-windows analysis to study the evolution of the obtained groups and we include the cluster dynamics and finally, section 5 present the conclusions and indicate directions for further research.

## 2. Data and methodology

Tourism ( $x$ ) is represented by international tourist arrivals per inhabitant (number of international arrivals divided by the number of inhabitants of the country), and Economic growth ( $y$ ) is represented by growth rate of per capita GDP. The data set includes 80 countries<sup>1</sup> over the period 1995–2016. Countries were selected trying to work with

<sup>1</sup> Albania, Argentina, Australia, Austria, Belgium, Bulgaria, Bolivia, Brazil, Botswana, Canada, Switzerland, Chile, China, Colombia, Costa Rica, Cyprus, Czech Republic, Germany, Dominican Republic, Algeria, Ecuador, Egypt, Spain, Estonia, Finland, France, United Kingdom, Greece, Guatemala, Croatia, Hungary, India, Ireland, Iceland, Israel, Italy, Jamaica, Jordan, Japan, Kenya, Kyrgyzstan, Cambodia, South Korea, Laos, Sri Lanka, Lithuania, Luxembourg, Latvia, Morocco, Mexico, Malta, Mauritius, Malawi, Malaysia, Namibia, Nicaragua, Netherlands, Norway, Nepal, Panama, Peru, Philippines, Poland, Puerto Rico, Portugal, Romania, Russia, Singapore, El Salvador, Slovakia, Slovenia, Sweden, Eswatini, Thailand, Tunisia, Tanzania, Ukraine, Uruguay, United States and South Africa.

countries from all geographical regions. Economic growth measured by per capita GDP (constant 2011US\$) was obtained from Maddison Project Database (2018). Tourism arrivals and population data are obtained from World Bank.<sup>2</sup>

The initial step to calculate the distance for subsequent cluster analysis is the symbolization of two-dimensional time series. Based on the analysis of symbolic time series the concept of “regime”, mainly used in studies on dynamic systems, is used. By transforming the original two-dimensional series into a symbolic series, while information is lost, we gain in understanding the dynamics that the time series follows. The transformation of two-dimensional time series into symbolic series is justified for different reasons. The most important one is that, when working with multi-dimensional time series, one cannot assume that the units of measure of the variables involved is the same (or at least, there is a functional relationship between them). This is what in the area of Econophysics is called “the speed along the axes”. (See Mantegna and Stanley, 1999) This prevents the use of Euclidean or similar metrics, since they start from the assumption that the units of measurement are the same. Therefore, symbolization is introduced as a way to employ a metric that allows to compare the dynamic trajectories of the different countries. Broadly speaking, symbolization permits us to focus on the trajectories countries follow over time. These trajectories imply changes from a phase, or regime, to another or, conversely, they stay in the same phase. Once we have defined these phases or regimes, we are able to describe and plot similar behaviors, which show the heterogeneity presented in the world economy regarding connections between tourism and economic growth.

We represent the dynamical economic performance of each country by the bi-dimensional time series of our variables: number of arrivals per inhabitant ( $x$ ) and growth rate of per capita GDP ( $y$ ).<sup>3</sup> To capture the dynamic evolution of both variables, we introduce the notion of economic regime. Each regime represents an economic situation that qualitatively contrasts from others. In this manner, we can describe the country’s dynamic as a sequence of economic regimes (see Brida, Puchet and Punzo, 2003; Brida & Punzo, 2003). The partition of the state space of tourism and growth rates of per capita GDP is defined by the annual averages of arrivals per inhabitant ( $\mu_x$ ), and growth rate of per capita GDP ( $\mu_y$ ). So we can divide the state space into the four regions determined by  $\mu_x$  and  $\mu_y$ .<sup>4</sup>

Fig. 1 shows a representation of this partition of state space, where the partition is determined by the average values of  $\mu_x$  and  $\mu_y$ . As we can expect, most data points (each data point representing the position of a country in the state space in a single year) lie around the average number of arrivals per inhabitant. In the same way, we observe that data points are evenly spread around the average growth rate. Lower part of Fig. 1 shows the attractor estimated by Kernel regression run with R 3.4.3 (RStudio 1.1.383). This attractor suggests the average dynamic trajectory of a country over time. Assuming that an average country goes in the kernel graph from left to right in the analyzed period, the transition over time would take a country from the left upper panel to the right upper panel. This path would imply that a typical country starts from a position where economic growth and tourism arrivals evolves

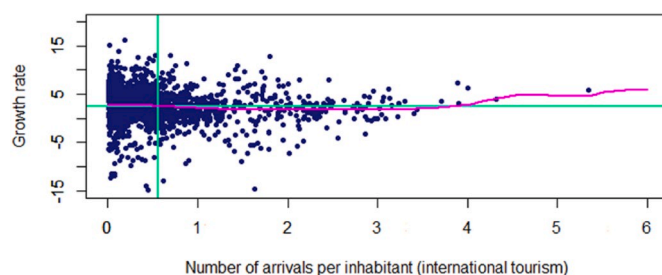


Fig. 1. Data partition in the state space for the set of 80 countries. Source: Maddison project database (MPD), World Bank and authors’ calculations.

similarly and when tourism arrivals are high enough economic growth is encouraged above the global mean. Therefore, this path suggests that high tourism, measured as number of per capita tourists, is related to higher economic growth over time but only from a certain point related to the volume of incoming tourist.

Our next step is to describe the qualitative behaviour of a country by using the notion of regime. In intuitive terms, an economic regime characterizes a particular qualitative conduct which is different from other dynamical behaviors. As previously mentioned, to capture this diversity of behaviors, we divide the state space into the four regions determined by the threshold values  $\mu_x$  and  $\mu_y$ . In the present case, the choice of thresholds is exogenous and, as a consequence, the results we obtain are contingent to these exogenous cut-offs. Future research could include the replication of this exercise for other convenient thresholds or other partition of regimes determined by different arguments (See Risso, 2018). We need to take into account that when the state space is divided into a large number of regions the statistical significance of some results can be affected. This is a consequence of having a finite time sample.

A change of regime of course signals some qualitative transformation. To explore these qualitative changes for every country, let us substitute a bi-dimensional time series  $\{(x_1, y_1), (x_2, y_2), \dots, (x_T, y_T)\}$ , by a sequence of symbols:  $s = \{s_1, s_2, \dots, s_T\}$  such that  $s_t = j$  if and only if  $(x_t, y_t)$  belongs to a selected state space region,  $R_j$ . We define four regions in the following way<sup>5</sup>

- $R_1 = \{(x, y) : x \leq \mu_x, y \leq \mu_y\}$
- $R_2 = \{(x, y) : x \leq \mu_x, y \geq \mu_y\}$
- $R_3 = \{(x, y) : x \geq \mu_x, y \geq \mu_y\}$
- $R_4 = \{(x, y) : x \geq \mu_x, y \leq \mu_y\}$

Regime 1 represents low level of both variables, economic growth and tourism (by “low” we mean below the average, and correspondingly by “high” we mean above the average). Regime 2 is characterized by low tourism but high economic growth, regime 3 by both high growth and tourism. Finally regime 4 shows low growth but high tourism. As showed by Figs. 2–4, very heterogeneous situations can be obtained. For instance, a large group of countries never lie in regimes 3 and 4 (e. g., Argentina, Australia, India, China, Chile). Conversely, a smaller group, only 16 countries, is never found in regimes 1 and 2 (e. g., France, Italy, Spain, Greece, Ireland). Finally, other group of countries, concretely 20 of them, have moved across all regimes (e. g., Canada, Uruguay, Latvia, Malaysia, Eswatini).

Figs. 2–4 show trajectories across regimes of selected countries. They reveal that there exists a large variety of performances.

Fig. 2 shows countries which have visited all the different regimes. For instance, Belgium, Canada, Jamaica and Uruguay exhibit an

<sup>2</sup> We use ‘International inbound tourists (overnight visitors)’. This measure considers the number of tourists who travel to a country, in which they do not have their usual residence, for a period not exceeding 12 months and the main purpose in visiting is an activity non-remunerated. The tourism data refer to the number of arrivals, not to the number of people traveling, if a person makes several trips to a country during a given period, is counted each time as a new arrival.

<sup>3</sup> Note that we are considering a variable in growth rate (picking up the dynamics) while the other is considered in levels (without considering the dynamics). In the future, the exercise could be carried out with both variables in levels or both variables in growth rates.

<sup>4</sup> For each year (in 1995–2016) we have  $\mu_x = \sum_{i=1}^{80} x_{iV}/80$  and  $\mu_y = \sum_{i=1}^{80} y_{iV}/80$

<sup>5</sup> The boundaries of the different regimes are defined by means of  $\geq$  or  $\leq$  because the probability of being in two regimes at the same time is 0.

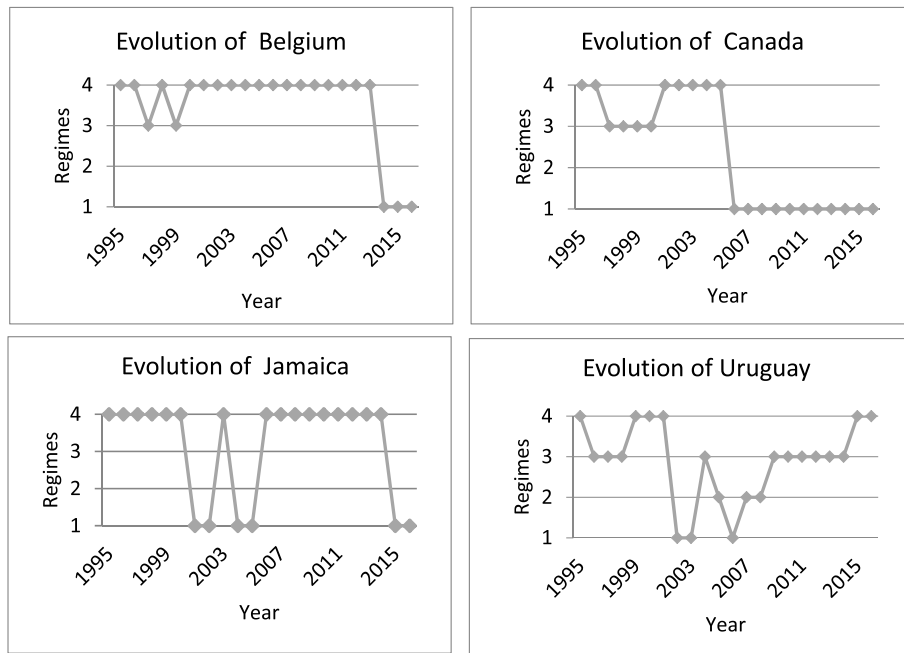


Fig. 2. Representation of regimes dynamics for Belgium, Canada, Jamaica and Uruguay for the period 1995–2016. Source: Maddison project database (MPD), World Bank and authors' calculations.

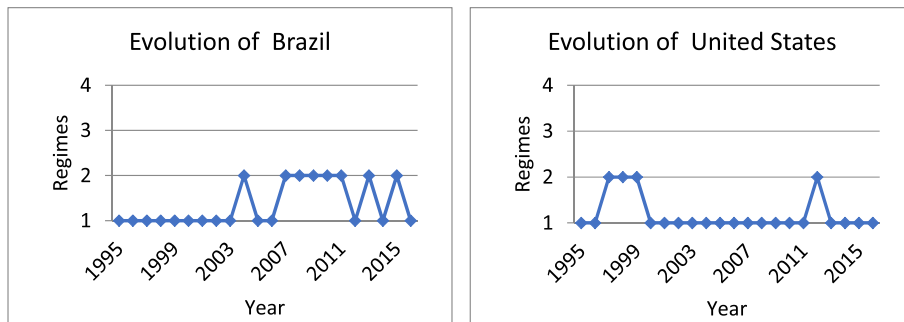


Fig. 3. Representation of regimes dynamics for Brazil and United States for the period 1995–2016. Source: Maddison project database (MPD), World Bank and authors' calculations.

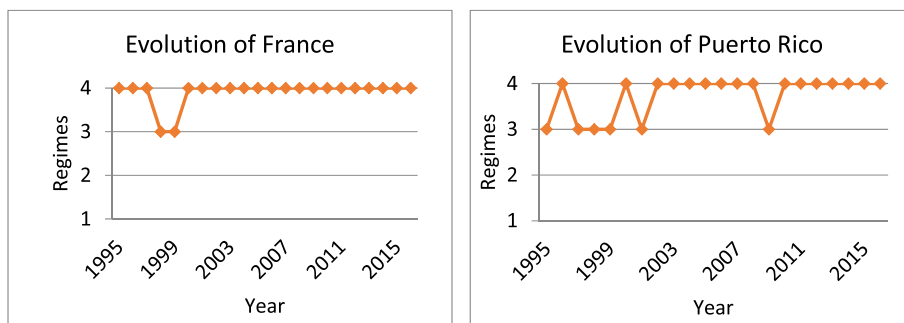


Fig. 4. Representation of regimes dynamics for France and Puerto Rico for the period 1995–2016. Source: Maddison project database (MPD), World Bank and authors' calculations.



irregular trajectory moving across regimes quite randomly over time. Fig. 3 selects countries where tourism is always behind the mean. In this case, United States stays most of the years behind the average growth rate while Brazil exhibits more variability regarding economic growth. Finally, Fig. 4 presents examples of countries with high levels of tourism. This case is similar to the one presented in Fig. 3: the more developed country remains in the low growth regime almost all years while Puerto Rico varies from low to high (in fact, is below or above the mean for that year). These figures reveal that heterogeneity in the tourism-growth dynamics is the rule rather than the exception. Table 1 (in Appendix) shows the percentage of time in each regime, for the set of 80 countries during the period 1995–2016, and additionally contains the codes to represent the different countries that are used in the next section.

**3. Minimum spanning tree (MST) and a hierarchical tree (HT)**

Heterogeneity seems to be therefore the rule in the economic growth-tourism trajectories of countries. To create a topology and hierarchy according to their economic-tourism performances we employ a non-parametric methodology based on the non-loop networks of Minimal Spanning Trees (MST) and Hierarchical Trees (HT), tools originally introduced in economics and finance by (Mantegna, 1999) and (Mantegna & Stanley, 2000). To obtain these trees we define a metric distance between the dynamical performance of every pair of countries. We define the “distance” between the paths of two countries measuring how close they are in their respective regime dynamics. To do this we introduce the notion of distance between symbolic sequences, where each sequence represents the way a country moves through regimes over time. Several distances can be postulated (see Piccardi, 2004; Molgedey & Ebeling, 2000; Tang, Tracy, Boozer, de Brown, & Brown, 1994, 1995 and 1997). We have chosen the notion of distance for symbolic time series most used: the binary distance. Given the symbolic sequences  $\{s_{it}\}_{t=1}^{t=T}$  and  $\{s_{jt}\}_{t=1}^{t=T}$  the distance between two countries,  $i$  and  $j$  is given by:

$$d(s_i, s_j) = \sum_{t=1}^{22} f(s_{it}, s_{jt}) : f(s_{it}, s_{jt}) = \begin{cases} 0 & \text{if } s_{it} = s_{jt} \\ 1 & \text{if } s_{it} \neq s_{jt} \end{cases}$$

that is, each of the 22 adding terms is 0 if countries  $i$  and  $j$  were in the same regime at that time, or 1 in case they are not in the same regime. Thus, we obtain a distance that takes value 0 in case both countries coincide in the very same regimes throughout the entire period, and it takes a maximum value of 22 in case they have not coincided at any time in the same regime during the considered period.

To build the Hierarchical Tree we employ the nearest neighbor single-linkage cluster algorithm as described in (Mantegna & Stanley, 2000). This technique uses an aggregative process, implying that in the first step the initial partition is formed considering each country as a cluster:  $A = \{A_1, A, \dots, A_n\}$ . The two nearest groups (those of less distance):  $A_i, A_j$  (with  $i = 1, \dots, n; j = 1, \dots, n; i \neq j$ ) are determined and they are grouped into a single cluster, forming the new partition:

$$A = \{A_1, A_2, \dots, A_i \cup A_j, \dots, A_n\}.$$

In the following stages, we continue grouping based on the minimum distances. In this process, the distance between clusters is given by the minimum distance between the individuals of each one, that is the distance between the clusters  $A_i$  (with  $n_i$  elements) and  $A_j$  (with  $n_j$  elements), is defined by:

$$d(A_i, A_j) = \text{Min}\{d(x_k, x_l)\}, \text{ with } x_k \in A_i, x_l \in A_j (k = 1, \dots, n_i; l = 1, \dots, n_j).$$

The Minimum Spanning Tree (Kruskal, 1956) is gradually constructed by connecting all the countries considering the minimum distance. This construction is represented in a graph of  $n$  vertices corresponding to each country and  $n-1$  links, where the most relevant links of each particular country are selected. In the first step, we take a

**Table 2**  
Shortest distances.

Link	Country 1	Country 2	Distance
1	Austria	France	0
2	Switzerland	Austria	1
3	China	Laos	1
4	Netherlands	Portugal	1
5	Italy	Switzerland	2

Source: Maddison project database (MPD), World Bank and authors' calculations

pair of individuals presenting the shortest distance, and we connect them. In the second step, we connect the pair of individuals with the second shortest distance, and we continue with this process until all countries are connected in a single tree.

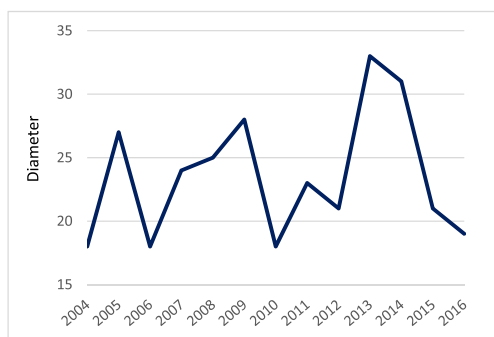
Table 2 shows the five shortest distances between all countries. Note that in the present exercise, the distance between Austria and France is zero, since these two countries have been lying in the same regime at each tick of the clock. Then, in the construction of the MST, the nodes representing Austria and France coincide. The next minimum distance is from Switzerland to Austria (or France), implying that Switzerland joins this first pair of countries. Subsequently, China and Laos form a new group of two countries. The next link involves another new pair of countries: Netherlands and Portugal. Subsequently, Italy links the group formed by Austria, France and Switzerland. Continuing in the same way, we construct the whole tree by adding one country at each step. As a result, the MST is built with 80 nodes and 79 links.

Figs. 5 and 6 shows the HT and the MST for the entire period for our group of 80 countries. A set of indicators were considered in order to determine the optimal number of clusters, using the Pseudo-F (Calinski & Harabasz, 1974) and the Pseudo- $t^2$  (Duda & Hart, 1973) methodologies. In the present exercise, both tests indicate that the optimal number of groups is two, and five countries that remain ungrouped (outliers). In particular, we find that Botswana, Canada, Eswatini, Latvia, Singapore, Tunisia and Uruguay to be “isolated” countries which do not belong to any of the two main clusters. The two clusters are composed by the following countries:

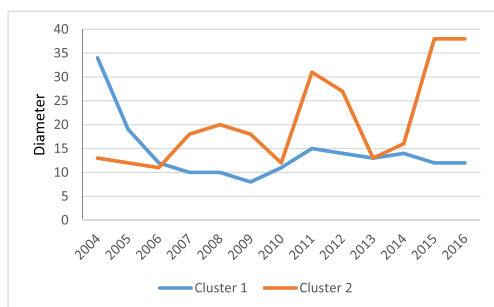
1. Cluster one is formed by Austria, Belgium, Bulgaria, Switzerland, Cyprus, Czech Republic, Spain, Estonia, France, Greece, Croatia, Ireland, Iceland, Italy, Jamaica, Luxembourg, Malta, Mauritius, Malaysia, Netherlands, Norway, Puerto Rico, Portugal and Slovenia. Countries in this group are characterized by trajectories through the state space in which they have stayed most of the times in regimes 3 and 4. As long as these regimes represent above the mean arrivals per inhabitant incoming tourism we call this cluster, *High Tourism Group*.
2. Cluster two is formed by Albania, Argentina, Australia, Bolivia, Brazil, Chile, China, Colombia, Costa Rica, Germany, Dominican Republic, Algeria, Ecuador, Egypt, Finland, United Kingdom, Guatemala, Hungary, India, Israel, Jordan, Japan, Kenya, Kyrgyzstan, Cambodia, South Korea, Laos, Sri Lanka, Lithuania, Morocco, Mexico, Malawi, Namibia, Nicaragua, Nepal, Panama, Peru, Philippines, Poland, Romania, Russia, El Salvador, Slovakia, Sweden, Thailand, Tanzania, Ukraine, Unites States and South Africa. This group shows the opposite behaviour than cluster one: their countries remain the whole period in regimes 1 and 2. As long as countries belonging this group remain most of the years below the tourist arrivals average, we have called it *Low Tourism Group*.

Note that inside the two groups one can detect the presence of geographical blocks, indicating that geographical closeness is relevant for the performance in economic growth and tourism. It would be possible to think that, from the tourism point of view, countries are perceived as being part of larger areas or regions where tourists decide to come. For example, inside group one we find Argentina and Chile or





**Fig. 7.** Evolution of the diameter of the MST for windows of 10 years. Source: Maddison project database (MPD), World Bank and authors' calculations.



**Fig. 8.** Evolution of the diameter of the MST for windows of 10 years. Source: Maddison project database (MPD), World Bank and authors' calculations.

in the dynamics for the whole group.

Taking into consideration that we identify two clear clusters with different dynamics, we further extend this analysis to our two different groups. Fig. 8 shows the evolution of the diameter –global distance– for each cluster in the same way than previously. When we look at our two groups, it is clear that the size of the clusters has experienced fairly important changes in their configuration. The figure shows that the low performance cluster two tends to span, while the high performance cluster 1 tends to be (more) compact. Consequently, countries in this group diverge in their performance over time. Clearly, the evolution of the global distance for the whole group of countries seems to be driven by the dynamics of Cluster 2. When comparing the composition of the clusters for the different time windows, it can be noted that certain countries move across clusters what can be interpreted as a change in the dynamical behaviour. Finally, taking the representative country of each cluster, we can analyse the evolution of the distance between the two groups. The empirical result shows that the distance between the two groups increases in time. That is, the dynamical behaviour of the groups tend to be more dissimilar. We can say that, as groups, a divergence process is observed. These results indicate that there are no fundamental traps to development of tourism, as countries particularly can change from one cluster to the other. In addition, this does not mean that growth is inevitable as predicted by the neoclassical theory as several countries always remain in the low-performance group, and the distance among countries inside the group tend to decrease.

## 5. Conclusions

Tourism is accepted to positively contribute to economic growth through very different and numerous direct and indirect channels. A growing academic literature has been developed during the last twenty years to deal with existing connections between economic growth and tourism (see Nunkoo et al., 2019). Contributions of this literature provide support for the TLGH and report positive and statistically significant estimates. However, most of the papers employ linear regressions, which by construction identify the same theoretical model for every possible empirical application. This study aims to contribute on the empirical discussion around the relationship between tourism and economic growth by employing a non-parametric, non-linear approach. In particular, we compare the qualitative dynamical behaviour of these two variables using the notion of economic regime and clustering tools based on the concept of minimum spanning trees and hierarchical trees (Mantegna, 1999). The dataset includes 80 developed and developing countries in the twenty-year period running from 1995 to 2016. Most of the findings in this paper have been eluded by the traditional analyses, which predicts inexorable growth towards a steady state. Note that the study examines the role at a different level of tourist arrivals and in this sense can be recognized as shedding the same light of those studies working on tourism-growth nexus using quantile regression analysis. Clearly quantile regression is closer to our methodology than traditional linear regression methods, as quantile regression is more robust in the presence of outliers or weaker linear correlation between variables. However, we do not look for the concept of correlation between the two main variables. In its place, our methods simply describe the paths of countries over time related to the regimes we define and, subsequently, we group these countries in clusters that observe a similar (and different) behaviour over time. In doing so, we identify and classify the heterogeneity observed in the growth-tourism nexus that arises from large panel data.

The non-parametric approach shows the existence of two clear clubs of countries that followed distinct sign patterns: dynamics of these two groups differs substantially. Note that the distribution of clusters shows that countries in a particular group are geographically close. (see Fig. 9 in the Appendix). The two main clusters that are identified over the whole time interval can be interpreted as two groups of countries with high and low performance in the tourism sector and are coherent with the business cycle. This result implies a heterogeneity in the relationship between economic growth and tourism. The evolution of their relative distance shows that the countries have tended to maintain the diameter of the whole group, but following an irregular path. Looking inside the two clusters we note that whereas low performance countries tend to diverge, the countries in the other cluster tend to have a more similar dynamical behaviour. When we study the evolution of the distance between an “average” High and Low Performance country we cannot reject the hypothesis that poor and rich countries have diverged. In sum, the dynamical behaviour of the clusters shows three main stylized facts: certain countries move across clusters; the low performance cluster tends to span, while the high performance one tends to be (more) compact; the distance between the two groups increases in time. All the results found (existence of clusters of countries, divergence/convergence between and intra-groups, etc.) are ex post, eliminating any selection bias. The empirical findings have important implications in providing policy-makers directions for achieving a path of growth that includes tourism as a key sector. Detecting a positive relationship between tourism and economic growth, also in terms of comparison with successful countries, is useful to governments that are prepared to develop tourism as a stimulus to their economy. In this sense, an additional and interesting policy implication can be addressed. As previously

noted, by using this methodology homogenous countries, in terms of the long-term growth-tourism relationship, emerge. This is especially useful for governments to find similar countries from which they can learn about their touristic policies. As the new structural economics (Lin, 2012) has highlighted it is important for countries to focus on improving their development policies looking at countries with better but similar performance trajectories and factor endowments.

The present study illustrates the need to further develop the validation of the TLGH not only with the use of innovative methodological approaches, for example, taking into account possible non-linearity between tourism and growth, but also revisiting the panel data exercises to validate the existence, or not, of homogeneity of the countries included in the studies. In this sense, most of the findings in this paper present a criticism of the traditional analyses of the TLGH, which predicted inexorable convergence toward a steady state by starting from linear modelling between the main variables. From a more dynamic and multidimensional perspective, this new approach has allowed us to uncover regularities and trends in economic behaviour. The results in this study contradicts the traditional analyses of convergence in the models of economic growth and tourism, which predict inevitable growth toward an equilibrium. From the proposed dynamic and multidimensional perspective, this method has allowed to discover regularities and trends in economic behaviour. We established the existence of clusters of performance, without having to condition the data a priori. In addition, we have showed that this clubs can present mobility and that the dynamic behaviour could tend to be more (or less) homogeneous. Note also that all the results found (existence of groups of countries with homogeneous dynamics, mobility between groups, etc.) are ex post, removing any selection bias.

## APPENDIX

**Table 1**  
Percentage of time in each regime, of the set of 80 countries for the period 1995–2016

N°	Code	Country	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	Total
1	ALB	Albania	5%	59%	18%	18%	100%
2	ARG	Argentina	50%	50%	0%	0%	100%
3	AUS	Australia	77%	23%	0%	0%	100%
4	AUT	Austria	0%	0%	9%	91%	100%
5	BEL	Belgium	14%	0%	9%	77%	100%
6	BGR	Bulgaria	14%	18%	41%	27%	100%
7	BOL	Bolivia	59%	41%	0%	0%	100%
8	BRA	Brazil	64%	36%	0%	0%	100%
9	BWA	Botswana	5%	18%	41%	36%	100%
10	CAN	Canada	50%	0%	18%	32%	100%
11	CHE	Switzerland	0%	0%	5%	95%	100%
12	CHL	Chile	45%	55%	0%	0%	100%
13	CHN	China	5%	95%	0%	0%	100%
14	COL	Colombia	59%	41%	0%	0%	100%
15	CRI	Costa Rica	50%	50%	0%	0%	100%
16	CYP	Cyprus	0%	0%	36%	64%	100%
17	CZE	Czech Republic	5%	9%	45%	41%	100%
18	DEU	Germany	86%	14%	0%	0%	100%
19	DOM	Dominican Republic	23%	77%	0%	0%	100%
20	DZA	Algeria	73%	27%	0%	0%	100%
21	ECU	Ecuador	59%	41%	0%	0%	100%
22	EGY	Egypt	55%	45%	0%	0%	100%
23	ESP	Spain	0%	0%	32%	68%	100%
24	EST	Estonia	0%	5%	68%	27%	100%
25	FIN	Finland	68%	32%	0%	0%	100%
26	FRA	France	0%	0%	9%	91%	100%
27	GBR	United Kingdom	77%	23%	0%	0%	100%
28	GRC	Greece	0%	0%	36%	64%	100%
29	GTM	Guatemala	82%	18%	0%	0%	100%
30	HRV	Croatia	0%	5%	59%	36%	100%
31	HUN	Hungary	41%	59%	0%	0%	100%
32	IND	India	14%	86%	0%	0%	100%

(continued on next page)

The proposed methodology allows to incorporate other variables into the analysis (economic, institutional, social, etc.), to compare the influence of such variables in the conformation of clubs up from changes in performance. This is matter of further research. In addition, it could be interesting to repeat the exercise using other variables representing tourism: tourism expenditure; per capita tourist arrivals, per capita tourism expenditure, etc. (see Rosselló – He, 2019). Finally, it is relevant to mention some limitations of this paper. For example, the time series data used in the study are not long, and the possibility that the measurements made to quantify the tourism sector could be not completely adequate.

### Author contribution notes

All the authors have contributed: Conceptualization, Methodology, Software, Data curation, Writing- Original draft preparation Investigation Writing- Reviewing and Editing

### Declaration of competing interest

None

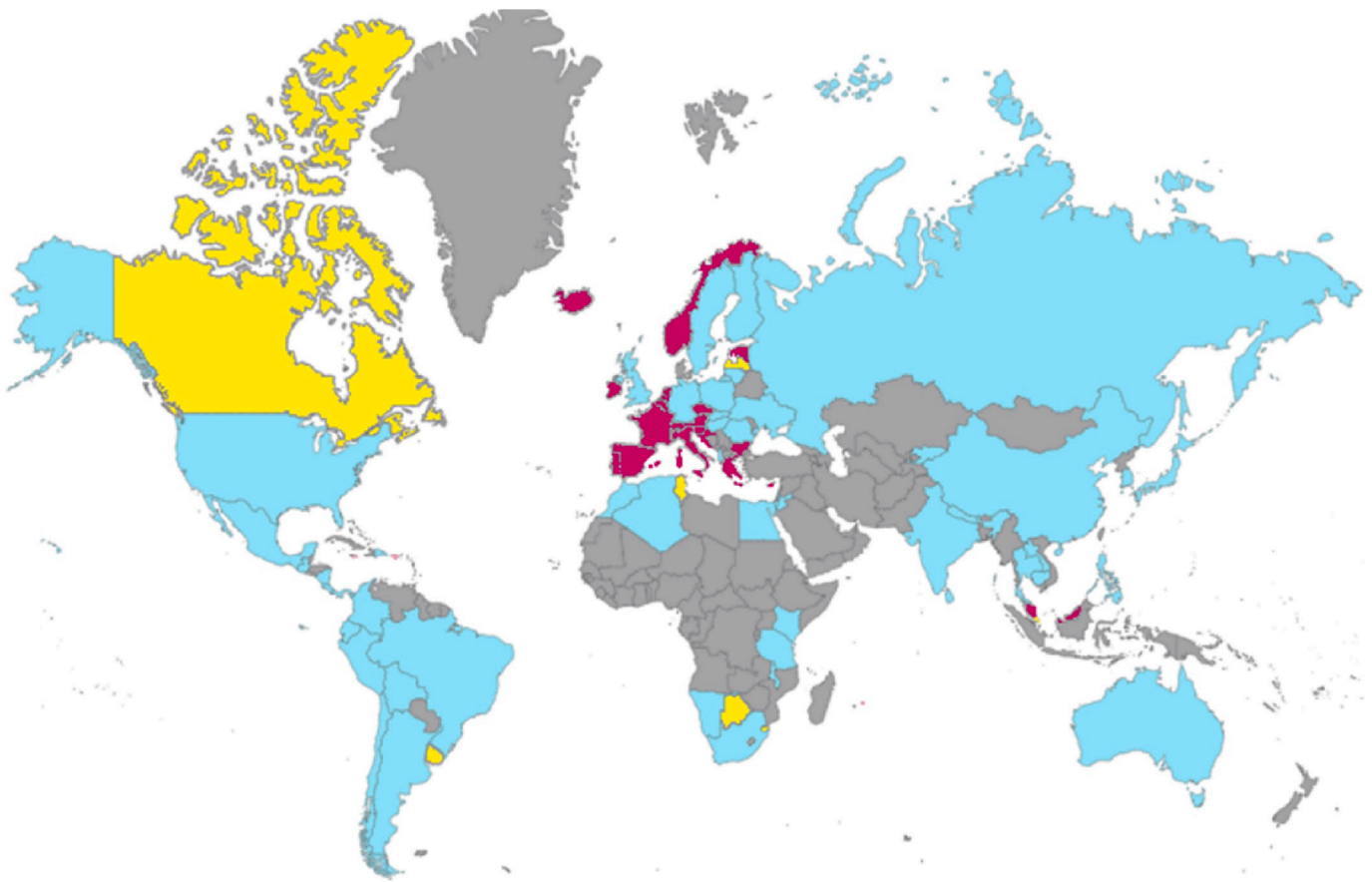
### Acknowledgements

Our research was supported by the UdelaR - CSIC, project GIDE. A preliminary version of this paper was presented at the 10th Workshop “Tourism: Economics and Management. Tourists as Consumers, Visitors and Travelers” In honor of Lionello F. Punzo. Siena (Italia), 30 May - 1 June 2019.



Table 1 (continued)

N°	Code	Country	R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	R <sub>4</sub>	Total
33	IRL	Ireland	0%	0%	55%	45%	100%
34	ISL	Iceland	0%	0%	50%	50%	100%
35	ISR	Israel	64%	32%	4%	0%	100%
36	ITA	Italy	0%	0%	5%	95%	100%
37	JAM	Jamaica	27%	0%	0%	73%	100%
38	JOR	Jordan	73%	27%	0%	0%	100%
39	JPN	Japan	86%	14%	0%	0%	100%
40	KEN	Kenya	64%	36%	0%	0%	100%
41	KGZ	Kyrgyzstan	36%	59%	5%	0%	100%
42	KHM	Cambodia	5%	95%	0%	0%	100%
43	KOR	South Korea	14%	86%	0%	0%	100%
44	LAO	Laos	9%	91%	0%	0%	100%
45	LKA	Sri Lanka	14%	86%	0%	0%	100%
46	LTU	Lithuania	9%	77%	14%	0%	100%
47	LUX	Luxembourg	0%	0%	41%	59%	100%
48	LVA	Latvia	5%	45%	36%	14%	100%
49	MAR	Morocco	50%	50%	0%	0%	100%
50	MEX	Mexico	64%	36%	0%	0%	100%
51	MLT	Malta	0%	0%	64%	36%	100%
52	MUS	Mauritius	0%	4%	73%	23%	100%
53	MWI	Malawi	64%	36%	0%	0%	100%
54	MYS	Malaysia	5%	27%	45%	23%	100%
55	NAM	Namibia	55%	45%	0%	0%	100%
56	NIC	Nicaragua	59%	41%	0%	0%	100%
57	NLD	Netherlands	0%	0%	23%	77%	100%
58	NOR	Norway	0%	0%	23%	77%	100%
59	NPL	Nepal	50%	50%	0%	0%	100%
60	PAN	Panama	32%	68%	0%	0%	100%
61	PER	Peru	36%	64%	0%	0%	100%
62	PHL	Philippines	55%	45%	0%	0%	100%
63	POL	Poland	14%	63%	23%	0%	100%
64	PRI	Puerto Rico	0%	0%	27%	73%	100%
65	PRT	Portugal	0%	0%	27%	73%	100%
66	ROU	Romania	36%	64%	0%	0%	100%
67	RUS	Russia	41%	59%	0%	0%	100%
68	SGP	Singapore	0%	0%	59%	41%	100%
69	SLV	El Salvador	77%	23%	0%	0%	100%
70	SVK	Slovakia	18%	82%	0%	0%	100%
71	SVN	Slovenia	0%	9%	64%	27%	100%
72	SWE	Sweden	64%	36%	0%	0%	100%
73	SWZ	Eswatini	50%	5%	18%	27%	100%
74	THA	Thailand	32%	68%	0%	0%	100%
75	TUN	Tunisia	23%	4%	50%	23%	100%
76	TZA	Tanzania	41%	59%	0%	0%	100%
77	UKR	Ukraine	50%	50%	0%	0%	100%
78	URY	Uruguay	14%	14%	45%	27%	100%
79	USA	United States	82%	18%	0%	0%	100%
80	ZAF	South Africa	82%	18%	0%	0%	100%



Note: Pink: high tourism group, sky blue: low tourism group, yellow: outliers

Fig. 9. Geographical distribution of the countries in each cluster

Source: Maddison project database (MPD), World Bank and authors' calculations.

Note: Pink: high tourism group, sky blue: low tourism group, yellow: outliers

## References

- Adamou, A., & Clerides, S. (2010). Prospects and limits of tourism led growth: The international evidence. *Review of Economic Analysis*, 3, 287–303.
- Balaguer, J., & Cantavella-Jordà, M. (2002). Tourism as a long-run economic growth factor: The Spanish case. *Applied Economics*, 34, 877–884.
- Bella, G. (2018). Estimating the tourism induced environmental Kuznets curve in France. *Journal of Sustainable Tourism*, 1–10.
- Brida, J. G. (2008). The dynamic regime concept in Economics. *International Journal of Economic Research*, 5(1), 55–76.
- Brida, J. G., Cortes-Jimenez, I., & Pulina, M. (2016a). Has the tourism-led growth hypothesis been validated? A literature review. *Current Issues in Tourism*, 19(5), 394–430.
- Brida, J. G., Lanzilotta, B., & Pizzolon, F. (2016b). Dynamic relationship between tourism and economic growth in MERCOSUR countries: A nonlinear approach based on asymmetric time series models. *Economics Bulletin*, 36(2), 879–894.
- Brida, J. G., Lanzilotta, B., Pereyra, J. S., & Pizzolon, F. (2015). A nonlinear approach to the tourism-led growth hypothesis: The case of the MERCOSUR. *Current Issues in Tourism*, 18(7), 647–666.
- Brida, J. G., PuchetAnyul, M., & Punzo, L. F. (2003). Coding economic dynamics to represent regime dynamics: A teach-yourself exercise. *Structural Change and Economic Dynamics*, 14, 133–157.
- Brida, J. G., & Punzo, L. F. (2003). Symbolic time series analysis and dynamic regimes. *Structural Change and Economic Dynamics*, 14, 159–183.
- Brida, J. G., Punzo, L. F., London, S., & Risso, W. A. (2011). Growth empirics: An alternative view to convergence. *Growth and change. A Journal of Urban and Regional Policy*, 42(3), 320–350.
- Calinski, R. B., & Harabasz, J. A. (1974). Dendrite method for cluster analysis. *Communications in Statistics - Theory and Methods*, 3(1), 1–27.
- Castro-Nuño, M., Molina-Toucedo, J. A., & Pablo-Romero, M. P. (2013). Tourism and GDP: A meta-analysis of panel data studies. *Journal of Travel Research*, 52(6), 745–758.
- Chiang, G. N., Sung, W. Y., & Lei, W. G. (2017). Regime-switching effect of tourism specialization on economic growth in asia pacific countries. *Economics*, 5(3), 23.
- Chingarande, A., & Saayman, A. (2018). Critical success factors for tourism-led growth. *International Journal of Tourism Research*, 20(6), 800–818.
- Chiu, Y. B., & Yeh, L. T. (2017). The threshold effects of the tourism-led growth hypothesis: Evidence from a cross-sectional model. *Journal of Travel Research*, 56(5), 625–637.
- Comerio, N., & Strozzi, F. (2019). Tourism and its economic impact: A literature review using bibliometric tools. *Tourism Economics*, 25(1), 109–131.
- Cristelli, M., Tacchella, A., & Pietronero, L. (2015). The heterogeneous dynamics of economic complexity. *PloS One*, 10(2), e0117174.
- De Vita, G., & Kyaw, K. S. (2017). Tourism specialization, absorptive capacity, and economic growth. *Journal of Travel Research*, 56(4), 423–435.
- Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. New York: Wiley.
- Eyuboglu, S., & Eyuboglu, K. (2019). Tourism development and economic growth: An asymmetric panel causality test. *Current Issues in Tourism*, 1–7.
- Fahimi, A., Saint Akadiri, S., Seraj, M., & Akadiri, A. C. (2018). Testing the role of tourism and human capital development in economic growth: A panel causality study of micro states. *Tourism Management Perspectives*, 28, 62–70.
- Fonseca, N., & Sánchez-Rivero, M. (2019a). Significance bias in the tourism-led growth literature. *Tourism Economics*, 1354816619833564.
- Fonseca, N., & Sánchez Rivero, M. (2019b). Granger causality between tourism and income: A meta-regression analysis. *Journal of Travel Research*, 0047287519851189.
- Gül, H., & Özer, M. (2018). Frequency domain causality analysis of tourism and economic activity in Turkey. *European Journal of Tourism Research*, 19, 86–97.

- Karimi, M. S. (2018). The linkage between tourism development and economic growth in Malaysia: A nonlinear approach. *International Economic Journal*, 32(1), 53–65.
- Kruskal, J. B. (1956). On the shortest spanning tree of a graph and the traveling salesman problem. *Proceedings of the American Mathematical Society*, 7, 48–50.
- Kumar, R. R., & Stauvermann, P. J. (2016). The linear and non-linear relationship between of tourism demand and output per worker: A study of Sri Lanka. *Tourism Management Perspectives*, 19, 109–120.
- Li, K. X., Jin, M., & Shi, W. (2018). Tourism as an important impetus to promoting economic growth: A critical review. *Tourism Management Perspectives*, 26, 135–142.
- Lin, J. Y. (2012). *New structural economics: A framework for rethinking development and policy*. Washington, DC: World Bank.
- Mantegna, R. N. (1999). Hierarchical structure in financial markets. *The European Physical Journal B*, 11, 193–197.
- Mantegna, R. N., & Stanley, H. E. (2000). *An introduction to Econophysics: Correlations and complexity in finance*. UK: Cambridge University Press.
- Matesanz, D., Torgler, B., & Ortega, G. J. (2013). Measuring global economic interdependence: A hierarchical network approach. *The World Economy*, 36(12), 1632–1648.
- Mayer, M., & Vogt, L. (2016). Economic effects of tourism and its influencing factors. *Zeitschrift fuer Tourismuswissenschaft*, 8(2), 169–198.
- Molgedey, L., & Ebeling, W. (2000). Local order, entropy and predictability of financial time series. *The European Physical Journal B*, 15, 733–737.
- Muslija, A., Satrovic, E., & Erbaş, C.Ü. (2017). Panel analysis of tourism-economic growth nexus. *Uluslararası Ekonomik Araştırmalar Dergisi*, 3(4).
- Nunkoo, R., Seetanah, B., Jaffur, Z. R. K., Moraghen, P. G. W., & Sannasee, R. V. (2019). Tourism and economic growth: A meta-regression analysis. *Journal of Travel Research*, 0047287519844833.
- Piccardi, C. (2004). On the control of chaotic systems via symbolic time series analysis. *Chaos*, 14(4), 1026–1034.
- Risso, W. A. (2017). Symbolic time series analysis and its application in social sciences. In *Time series analysis and applications*. IntechOpen.
- Risso, W. A. (2018). Tourism and economic growth: A worldwide study. *Tourism Analysis*, 23(1), 123–135.
- Rosselló-Nadal, J., & He, J. (2019). Tourist arrivals versus tourist expenditures in modelling tourism demand. *Tourism Economics*, 1–16.
- Seetanah, B., Nunkoo, R., Sannasee, R. V., Georges, P., & Jaffur, W. M. Z. R. K. (2017). A meta-analysis of the tourism and economic growth nexus. *BEST EN Think Tank XVII: Innovation and Progress in Sustainable Tourism*, 180.
- Škrinjarić, T. (2019). Examining the causal relationship between tourism and economic growth: Spillover index approach for selected CEE and SEE countries. *Economics*, 7(1), 19.
- Sokhanvar, A., Çiftçioğlu, S., & Javid, E. (2018). Another look at tourism-economic development nexus. *Tourism Management Perspectives*, 26, 97–106.
- Tang, C. F., & Tan, E. C. (2015). Tourism-led growth hypothesis in Malaysia: Evidence based upon regime shift cointegration and time-varying Granger causality techniques. *Asia Pacific Journal of Tourism Research*, 20(sup1), 1430–1450.
- Tang, C. F., & Tan, E. C. (2018). Tourism-led growth hypothesis: A new global evidence. *Cornell Hospitality Quarterly*, 59(3), 304–311.
- Tang, X. Z., Tracy, E. R., Boozer, A. D., de Brauw, A., & Brown, R. (1995). Symbol sequence statistics in noisy chaotic signal reconstruction. *Physical Review E*, 51(5), 3871–3889.
- Tang, X. Z., Tracy, E. R., Boozer, A. D., de Brown, A., & Brown, R. (1994). Reconstruction of chaotic signal using symbolic data. *Physics Letters A*, 190, 393–398.
- Tang, X. Z., Tracy, E. R., & Brown, R. (1997). Symbol statistics and spatio-temporal systems. *Physica D*, 102, 253–261.

- Wu, P. C., Liu, S. Y., Hsiao, J. M., & Huang, T. Y. (2016). Nonlinear and time-varying growth-tourism causality. *Annals of Tourism Research*, 59, 45–59.
- Zhang, J., & Cheng, L. (2019). Threshold effect of tourism development on economic growth following a disaster shock: Evidence from the wenchuan earthquake, PR China. *Sustainability*, 11(2), 371.



Juan Gabriel Brida is Full Professor of Economic Dynamics at the Faculty of Economics - Universidad de la República, Uruguay. He received his PhD in Economics from the University of Siena, in Italy. His research interests and expertise are in the areas of tourism economics and economic dynamics.



David Matesanz is Professor of the Applied Economic Department - Universidad de Oviedo, Spain. He received his PhD in Economics from the Universidad Autonoma de Madrid. His research interests and expertise are in the areas of international economics, tourism economics and complex systems.



Verónica Segarra is a PhD student in Economics (Universidad de Murcia, Spain). He received a Master in Economics (FCEA - Universidad de la República, Uruguay) and a Bachelor in Economics (FCEA - Universidad de la República, Uruguay). He is Assistant Professor of the Department of Economics (FCEA - Universidad de la República, Uruguay). Member of the Research Group in Economic Dynamics (GIDE) - Department of Mathematical-Quantitative Methods (FCEA - UdelaR, Uruguay). Areas of knowledge and interest are economic dynamics, tourism economics and complex systems.