Publisher's notice

This *Tourism Economics* Fast Track paper has been peer-reviewed, revised and fully accepted for publication in *Tourism Economics*. However, this is an **unedited manuscript** and will undergo a rigorous editing process before its appearance in an issue of the journal. This means that this manuscript version of the paper may not conform to journal style in terms of presentation, spelling and other usages. It may also contain minor errors of typography, grammar, spelling, referencing, etc, all of which will be corrected in the processes of copy-editing and proofreading.

Tourism Economics operates a Fast Track online publication system so that papers can be published and made available almost immediately on final acceptance by the journal.

Citing this article

Each *Tourism Economics* Fast Track article is given a DOI. When the paper is assigned to an issue, this DOI will automatically be transferred to the article in the journal issue.

This version of the article may be cited using the DOI. Citations should include the author's or authors' name(s), the title of the article, the title of the journal followed by the words 'Fast Track', the year of Fast Track publication and the DOI. For example:

Smith, J. (2016), 'Article title', *Tourism Economics* Fast Track, DOI xxxxxxxx.

Once the paper has been published in an issue of the journal, the DOI will automatically resolve to that final version and the article can be cited in accordance with normal bibliographical conventions.

doi: 10.5367/te.2015.0536

RESEARCH NOTE: ECONOMIC CRISES AND MARKET PERFORMANCE – A MACHINE LEARNING APPROACH

José Francisco Perles-Ribes (corresponding author)

(jose.perles@ua.es)

Ana Belén Ramón-Rodríguez

(anar@ua.es)

Luis Moreno-Izquierdo

(luis.moreno@ua.es)

Martín Sevilla-Jiménez

(martin.sevilla@ua.es)

Department of Applied Economic Analysis

Faculty of Economics and Business Sciences

University of Alicante

Campus San Vicente del Raspeig

03080 Alicante, Spain

Abstract:

This note analyses the relationship between economic crises and tourism performance in Spain during the period 1970–2013 using machine learning techniques. Specifically, a regression tree is estimated to confirm that, although the dynamics of Spanish tourism performance is influenced by the general variables established by the literature, the crisis periods disrupt the natural functioning of these dynamics, provoking disturbances that affect the tourism market position of destinations to a greater extent than expected. Conversely to other econometric techniques, machine learning approach allows us to achieve greater flexibility and enriches the information estimating the inter-relations and thresholds operating in this context.

Keywords: Market performance, Machine Learning, Regression Trees, Spain

The relationship between economic crises and tourism competitiveness is a very promising field of research. Recently, three papers by Perles and Ramón (2013), Perles, Ramón, Rubia and Moreno (2013) and Perles, Ramón, Sevilla and Rubia (2014) have attempted to fill this gap in the existing literature, thereby opening a debate on this subject. Perles, Ramón, Rubia and Moreno (2013) analyzed the long-term implications of economic crises for Spain's tourism performance, using market share as a proxy for competitiveness and the unit root test to determine the persistence of the effects of economic crises on tourism destinations. The authors concluded that studies undertaken from a competitiveness perspective enrich analysis based solely on a demand interpretation. Meanwhile, using vector autoregression (VAR) techniques and the Granger causality approach, Perles and Ramón (2013) explored the differential effects that economic crises generated in tourism destinations, depending of the destinations' mature or emerging status. Finally, Perles, Ramón, Sevilla and Rubia (2014) provides the theoretical foundation of the model and estimate a threshold model for Spain's tourism performance during the period 1970-2013.

This note goes beyond attempting to estimate a regression tree for the same dataset used by Perles, Ramón, Sevilla and Rubia (2014). Like the latter study a non-linear approach lets us to achieve greater flexibility and enriches the information estimating the interrelations and thresholds operating in this context. In any case, the most innovative aspect of this note is the use of machine learning techniques to deal with tourism destinations competiveness issues.

But previously, a note of caution on the use of market share as proxy of tourism competitiveness is needed. As in previous papers of Perles and Ramón (2013), Perles, Ramón, Rubia and Moreno (2013) and Perles, Ramón, Sevilla and Rubia (2014), in this note market share (in terms of visitor arrivals) is not considered a true indicator of the competitiveness of a tourist destination but rather an indicator of international tourism success.

Competitiveness and market share are not the same. Competitiveness gets its justification as an antecedent of tourism performance or success where market share represents a most relevant indicator. However, Perles, Ramón and Sevilla (2014) show the usefulness and limitations of market shares as a proxy of the competitiveness on tourism destinations and justify its use at least in historical empirical analysis where other indicators are not available.

Of course, this use in empirical analysis cannot serve to justify the implementation of tourism policies focused exclusively on objectives of growing market shares from a practical point of view. Nowadays, the literature agrees that policy should seek to achieve a balanced, inclusive and sustainable development of tourist destinations. And this ultimate goal cannot be achieved with only a policy focused exclusively on market share (Perles, Ramón and Sevilla, 2014).

Leaving aside these terminological considerations and with regard to the econometric techniques, in this paper we use, as explained above, Classification and Regression Trees (CART). Following Loh (2008) classification and regression trees are machine-learning methods for constructing prediction models from data. The models are obtained by recursively partitioning the data space and fitting a simple prediction model within each partition. As a result, the partitioning can be represented graphically as a decision tree.

Building a decision tree for a Y a response variable or class from inputs $X_1, X_2, \ldots X_p$, the process of growing a binary tree involves proposing many possible data cuts and then choosing best cuts based on simultaneous competing criteria of predictive power, cross-validation strength, and interaction with other chosen cuts. The split which maximizes the reduction in impurity is chosen, the data set split and the process repeated. Splitting continues until the terminal nodes are too small or too few to be split.

The classic CART algorithm was popularized by Breiman, Friedman, Olshen, & Stone (1984) and Ripley (1996). Since then, other algorithms have been developed (see Loh, 2008 for a review) to increase computational efficiency. Here, we use the *'tree'* package (Ripley, 2014) and R 3.1.2 language programming (R Core Team, 2014) to estimate two regression trees where the variation of Spain's market share is the dependent variable. Both models considers as explanatory variables variation of GDP, the variation of international price competitiveness adjusted by exchange rates (RCPI), the variation Gross Capital Formation (GKF) and the variation of cement consumption (representing generic investment), the variation of inward and outward Foreign Direct Investment (FDI) and the variation of Spanish unemployment rate (as expectations). But the models differ in the lags specification of our set of explanatory variables as explained below.

Table 1 lists the variables considered. Like the previous study conducted by Perles, Ramón, Sevilla and Rubia (2014), here the variables are transformed into their logarithmic difference and should be interpreted in terms of variation rates.

Table 1 about here

Figure 1 show the short-term tree due that all variables appear without lags. According to this figure the most relevant determinant of tourism success in each destination is their relative price competiveness (ldRCPI). Losses of 5 percent or more in this competitiveness lead to years with decreases in market share of 9 per cent. A second group of variables determining destination's success is represented by national investment (ldCEMENT and ldBEDS) also representing the supply growth rate of destinations. A growth of 0.23 per cent in cement consumption and 0.82 per cent in beds capacity is usually associated with increases in Spain's tourism market share from 1 to 6 per cent depending on behavior of relative price competitiveness (ldRCPI) and international demand (ldGDPUK). Conversely, decreases in generic investment are associated with decreases in market share depending on the Spanish economic cycle situation (ldGDP).

Figure 1 about here

In order to be consistent with the dynamics of the mechanism described by Perles, Ramón, Sevilla and Rubia (2014), in the Figure 2 the variables representing the generic investment of the supply channel are introduced with two (GKF or CEMENT) or three lags (both types of FDI). Meanwhile, the variables representing the demand channel and the business cycle are included with only one lag. The introduction of these lags also prevents problems of endogeneity. Only price competitiveness (RCPI) is included in our regressions without lags, because the effect of this variable on competitiveness can be considered as almost immediate. This tree represents the mid or long-term mechanisms.

Figure 2 about here

The most important finding is that price competitiveness remains the most relevant transmission mechanism. This result should not be surprising since this variable appears in the model without lags. Conversely, the main difference observed over the previous tree affects the foreign direct investment (ldFDIINW). Negative growth rates of inward foreign direct investment of 12 per cent or more are associated with reductions of market share of 5 per cent. Also an important difference over the short-term tree is related with BEDS variable. If Spain's GDP is growing below 1.9 percent, decreases in beds capacity of 1.4 per cent could be associated with market share reductions of 1 percent. Otherwise, decreases in beds capacity leas to an increase of market share from 2 to 7 per cent.

Regarding the variable of interest –the role of the economic cycle- both models reflect that during the crisis period the loss of market share is more relevant (bigger coefficients) than in the expansion phase. Thus, the relevance of economic crises for tourism competitiveness is also confirmed. The advance from previous studies is that in this paper the impact of the crises is considered with respect to other determinants of tourism destinations competitiveness.

The obtained results confirm that the dynamics of Spanish tourism performance is influenced by the general variables established by the literature, namely price competitiveness and generic and tourism investment determinants of this competitiveness. However, the crisis periods disrupt the natural functioning of these dynamics provoking disturbances in the determinants of competitiveness that affect the tourism market position of destinations to a greater extent than expected.

Although the study of the consequences of crises on tourism is not a new research field, the original aspect of this paper is that uses a machine learning approach, which is an innovative aspect of this paper. Further research in line with this study should improve the estimations by reducing the variance prediction through bootstrap aggregation or other techniques.

REFERENCES:

Breiman, L., Friedman, J., Olshen, R., & Stone, C. (1984). Classfication and Regression Trees. Belmont. California. Wadsworth.

Loh, W. Y. 2008. Classification and regression tree methods. Pages. 315–323. in F. Ruggeri, R. S. Kenett, F. W. Faltin (editors). *Encyclopedia of statistics in quality and reliability*. Wiley Chichester, UK.

Perles, J and, Ramón, A. (2013). Economic cycles, asymmetric crises and tourism competitiveness: emerging versus mature destinations. *Economics Bulletin*, 33 (3): 2080-2097

Perles, J., Ramón, A., Rubia, A. and Moreno, L. (2013). Economic crisis and tourism competitiveness in Spain: permanent effects or transitory shocks? *Current Issues in Tourism*, in press. DOI 10.1080/13683500.2013.849666

Perles, J., Ramón, A. and Sevilla, M. (2014). La cuota de mercado como indicador de competitividad en los destinos turísticos: sentido y limitaciones. Cuadernos de Turismo, 34: 265-285.

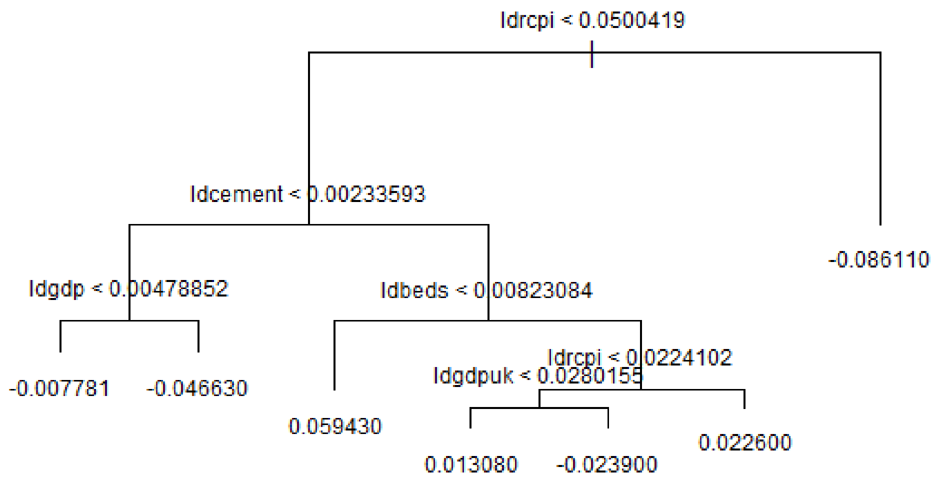
Perles, J., Ramón, A., Sevilla, M. and Rubia, A. (2014). The effects of economic crises on tourism success: an integrated model. *Tourism Economics*, in press http://dx.doi.org/10.5367/te.2014.0428

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ripley, B. (1996). Pattern Recognition and Neural Networks. Cambridge, England: Cambridge University Press.

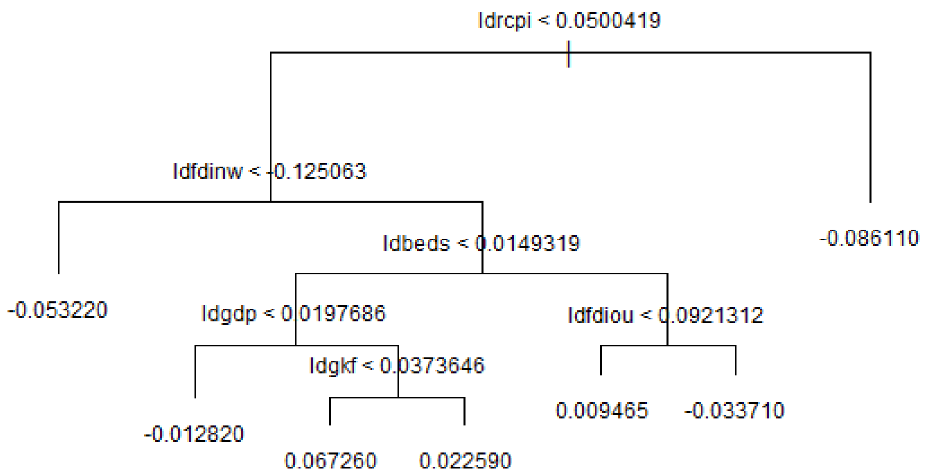
Ripley, B. (2014). tree: Classification and regression trees. R package version 1.0-35.

Figure 1: Contemporary regression tree (no lagged variables)



Author's own elaboration

Figure 2: Mid-term regression tree (tourism investment lagged variables)



Author's own elaboration

TABLES AND FIGURES

Table 1 Empirical analyses, variables used and sources.

Mechanism / Variable	Variable	Source / Observations
Dependent variable	Spain’s international tourism market share (logarithmic difference) (MSHARE)	Institute of Tourism Studies and WTO
Independent variables		
Domestic economic cycle	Real GDP of Spain (logarithmic difference) (GDPESP)	OECD
Foreign economic cycle	Real GDP of United Kingdom (logarithmic difference) (GDPESP)	OECD
Demand Mechanism		
Spain’s competitiveness index	Real effective exchange rate (RCPI)	OECD
Spain’s Demand Expectations	Unemployment rate (UNEMPLOY)	OECD
Supply Mechanism		
Tourism investment	Hotel beds in Spain (logarithmic difference) (BEDS)	General Secretariat of Tourism and the Spanish National Statistics Institute (INE)
Foreign Direct Investment	Inflows of FDI in nominal US dollars (logarithmic difference) (FDIin)	UNCTAD
	Outflows of FDI from Spain in nominal US dollars (logarithmic difference) (FDIout)	UNCTAD
Generic investment	Gross Fixed Capital Formation (logarithmic difference) (GKF)	OECD
	Spanish Cement Consumption (logarithmic difference) (CEMENT)	OFICEMEN

Author’s own elaboration