European Research Studies Journal
Volume XXIX, Issue 1, 2026

pp. 332-355

Short-Term Disruptions and Recovery Patterns in Spanish Hotel Activity: Insights from Quantitative and Qualitative Evidence

Submitted 23/11/25, 1st revision 05/12/25, 2nd revision 20/12/26, accepted 05/02/26

Guglielmo Maria Caporale[1], Luis Alberiko Gil-Alana[2], Carlos Poza[3], José L. Ruiz-Alba[4]

Abstract:

Purpose: *This paper examines the impact of Covid-19 pandemic on Spanish hotel activity to establish whether it has had temporary or permanent effects.*
Project/methodology/approach: *The analysis is based on both quantitative and qualitative approaches. For the former, data from Google Trends and the Spanish Statistical Office (INE) are collected to create a Leading Hotel Activity (LHA) index and fractional integration and cointegration methods are applied. For the latter, online interviews of a focus group in the Spanish hotel sector are conducted. The analysis also distinguishes between the five main source countries for Spain and the main five tourists regions in Spain.*
Findings: *The results show that the impact of Covid-19 shock on Spanish hotel activity was temporary, and that it disappeared at a faster rate in the case of the Balearic Islands and of tourists from Germany. The qualitative evidence indicates a strong linkage between intentions and behaviour in the Spanish tourism sector.*
Practical implications: *The findings indicate that the effects of Covid-19 on Spanish hotel activity were temporary. The Leading Hotel Activity (LHA) index based on Google Trends emerges as a useful tool for anticipating demand and supporting managerial and destination-level planning. Moreover, differences in recovery across regions and source markets underline the importance of targeted strategies, while the strong link between online search intentions and hotel stays highlights the role of confidence-building measures in accelerating demand recovery.*
Originality/value: *This paper combines fractional integration methods with qualitative evidence to analyse the persistence of shocks in Spanish hotel activity. It introduces a novel Leading Hotel Activity (LHA) index based on Google Trends and provides new evidence on the link between online search intentions and hotel stays across source markets and destination regions.*

Keywords: *Tourism, Spain, fractional integration, Leading Hotel Activity (LHA) index.*

[1]*Brunel University of London, London, UK.*
[2]*Corresponding author, University of Navarra, Pamplona, Spain and Universidad Francisco de Vitoria, Madrid, Spain, e-mail: alana@unav.es;*
[3]*Universidad Francisco de Vitoria, Madrid, Spain.*
[4]*University of Westminster, London, UK.*

JEL codes: *Z30, C22.*

Paper type: *Research article.*

Acknowledgement: *Luis A. Gil-Alana gratefully acknowledges financial support from the Grant D2023-149516NB-I00 funded by MCIN/AEI/ 10.13039/501100011033, and also from an internal project of the Universidad Francisco de Vitoria. A working paper version of this manuscript can be found at Caporale et al. (2025).*

1. Introduction

Tourism sector activity is often analysed using big data sets based on Google Trends and Baidu (Choi and Varian, 2012; Varian, 2014; Bangwayo and Skeete, 2015; Li *et al.,* 2017; Sun *et al.,* 2019). Search engine data can also be used to forecast tourist numbers (Padhi and Pati, 2017).

It is well known that the Covid-19 pandemic deeply affected the tourism and hospitality industry (Rivera, 2020), with travel restrictions showing its vulnerability (Kaushal and Srivastava, 2021). Investigating how tourist searches were influenced by it could help to produce better forecasts of tourist numbers.

This paper aims to provide new evidence on the impact of the Covid-19 pandemic on the Spanish tourism sector, and in particular on the hotel business. For this purpose, first we create a leading hotel activity (LHA) index by using Google Trends keywords; next we compare the results with data on foreigners' overnight stays in Spanish hotels from the Spanish Statistical Office (INE); finally, we apply both qualitative and quantitative methods to assess the effects of the Covid-19 pandemic on Spanish hotel activity.

The former consist of advanced time series methods, specifically fractional integration tests allowing to establish whether the effects of shocks such as the Covid-19 pandemic are transitory or permanent; the latter involve open-ended interviews with Spanish hotel industry experts. Other issues investigated in the paper are whether or not hotel managers can accurately forecast foreigners' overnight stay trends, and how many months in advance they can predict hotel demand behaviour.

For these purposes, first the usefulness of the leading hotel activity index (LHA) for forecasting foreign residents' overnight stays is assessed; then the relative importance of the five main leading source markets for Spain (the UK, Germany, France, Italy and the US) and of the principal five tourist destination regions in Spain (Malaga, Tenerife, Ibiza, Madrid and Barcelona) are analysed.

The layout of the paper is the following: Section 2 briefly reviews the most relevant empirical studies; Section 3 outlines the methodology; Section 4 describes the data and discusses the main empirical findings; Section 5 offers some concluding remarks.

2. Literature Review

2.1 Big Data, Leading Indicators and Tourism

The increasing accessibility of Big Data has transformed the landscape of economic forecasting by enabling earlier and more informed decision-making (Szármes, 2015). Within this framework, numerous scholars have applied text mining techniques to anticipate macroeconomic fluctuations (Antenucci *et al.,* 2014; Huang *et al.,* 2018; Poza and Monge, 2020; Varian, 2014), as well as to examine mesoeconomic developments in sectors such as tourism and the automobile industry (Bangwayo and Skeete, 2015; Choi and Varian, 2012).

Among these contributions, Poza and Monge (2020) proposed a novel composite index based on Google Trends search activity. Their approach integrated fractional cointegration analysis and continuous wavelet transformation to evaluate the extent to which the index could serve as a leading indicator of gross domestic product (GDP) trends.

Social media data has also proved instrumental in detecting economic shifts. Antenucci *et al.* (2014), for example, analysed millions of tweets to detect labour market disruptions, constructing a real-time proxy for job losses that aligned closely with official unemployment claims. Similarly, Dong *et al.* (2017) highlighted the value of geo-referenced digital data in capturing immediate economic dynamics.

Search engine data has likewise been found to provide relevant information for behavioural and financial forecasting. Varian (2014) demonstrated that online search trends are effective in predicting consumption and market behaviour. Toole *et al.* (2015) examined mobile phone records across Europe to measure employment shocks, while Pappalardo *et al.* (2016) introduced a framework for nowcasting socio-economic conditions using anonymised mobile mobility data.

Other researchers have focused on sentiment and media analysis as a means of extracting macroeconomic signals. Huang *et al.* (2018) evaluated the use of news sources to assess economic sentiment. Bernanke (2008) emphasized that public confidence indices play a crucial role in influencing economic performance, and Koenig (2002) argued that such indices may provide more timely insights than traditional lagging indicators.

There is also a growing body of work that applies these insights to financial and labour market contexts. Bollen *et al.* (2011) used sentiment data from Twitter to forecast equity market fluctuations. D'Amuri and Marcucci (2017) predicted U.S. unemployment using Google search data, while Hisano *et al.* (2013) employed news content to model financial market volatility.

In a related effort, Choi and Varian (2012) showed how search engine data can predict shifts in automotive sales, labour markets, and tourism. Likewise, Scott and Varian (2012) found that specific queries—such as "vehicle shopping" or "file for unemployment"—function as real-time economic indicators.

In the tourism sector, online behaviour has been directly linked to travel demand. Bangwayo and Skeete (2015) created a tourism leading indicator using Google search data for terms like "hotels" and "flights" from various origin markets to Caribbean destinations. Their study applied an Autoregressive Mixed-Data Sampling (AR-MIDAS) model and concluded that Google Trends improves forecasting accuracy in the tourism domain.

Further evidence is provided by Yang *et al.* (2015), who identified cointegration between online search queries (from both Google and Baidu) and visitor arrivals in Hainan, China. Their results showed that including such data improves forecast precision in autoregressive models. Similarly, Önder and Gunter (2016) found that integrating Google Trends data enhanced the accuracy of tourism demand forecasts for Vienna when compared to traditional models.

Rivera (2016) used search query volume (SQV) data from Google Trends to anticipate hotel registrations in Puerto Rico. Sun *et al.* (2019) employed machine learning techniques combined with Baidu and Google search data to forecast tourist arrivals in China. Li *et al.* (2017), meanwhile, developed a generalized dynamic factor model based on a composite search index to project tourism demand in Beijing, showing this method to be superior to traditional principal component analysis in terms of forecasting performance.

2.2 The Theory of Planned Behaviour (TPB) and Google Trends Searches in Tourism

The literature on tourist destination planning and search behaviour draws on four major theoretical paradigms (Erawan *et al.,* 2011); 1) the motivational approach, which is related to the needs of a person; 2) the economic approach, which assumes that consumers are rational and apply cost–benefit criteria when looking for data; 3) the consumer information processing approach, according to which consumers search for data to achieve their purposes; 4) the Theory of Planned Behaviour (TPB), which connects one's beliefs and behaviour and argues that attitude and subject norms drive the individual's intentions and behaviours.

Kim *et al.* (2016) noted that the TPB framework suggests that factors such as attitude and behavioural control are associated with tourist behaviour and intention for online search. Further, Padhi and Pati (2017) pointed out that information processing has been identified as one of the best predictors of tourist behaviour in selecting travel destinations through online keyword searches.

Some researchers have used the TBP approach to analyse specific types of behaviour, such as consumers' intentions to visit a green hotel (Han *et al.,* 2010), tourists' intentions to visit a country (Quintal *et al.,* 2010) or travellers' intentions to choose a long/short distance holiday destination (Bianchi *et al.,* 2017).

Their findings suggest that intentions may not be the only foundation of behaviour, but they are significant enough to forecast it.

As previously mentioned, recent studies concerning tourist behaviour have analysed tourism destination selection using Google Trends data, and have revealed many aspects of tourists' keyword-based queries.

Specifically, it appears that keyword-based online searches are closely linked to actually visiting a place (Anton and Lawrence, 2016; Li *et al.,* 2017) also, anticipated emotions can play an important role in the decision-making process (Kim *et al.,* 2013) in light of the TPB, and fear is an emotion that can influence travelling decisions affected by the risk of being infected by Covid-19, as tourism is a complicated psychological process (Cutler and Carmichael, 2010).

The present study draws from both these two theories, namely the TPB (Ajzen, 1991) and the consumer information processing approach (Bettman, 1979; Chung *et al.,* 2015), to interpret the empirical findings and gain better insights into tourist behaviour in the selection of destinations.

3. Research Methodology

We use a mixed methods approach combining quantitative and qualitative techniques. Specifically, we follow an explanatory sequential model (Creswell, 2013), starting with a quantitative study using secondary data and then studying primary data obtained by a qualitative online focus group to gain a deeper understanding of the results from the quantitative investigation.

For the quantitative analysis we use fractional integration and cointegration methods which shed light on the degree of persistence of the series, the speed of adjustment towards the long-run equilibrium, and the transitory or permanent effects of shocks. We collect data from two sources, namely Google Trends, as in Havranek and Zeynalov (2021), and the Spanish Statistical Office (INE).

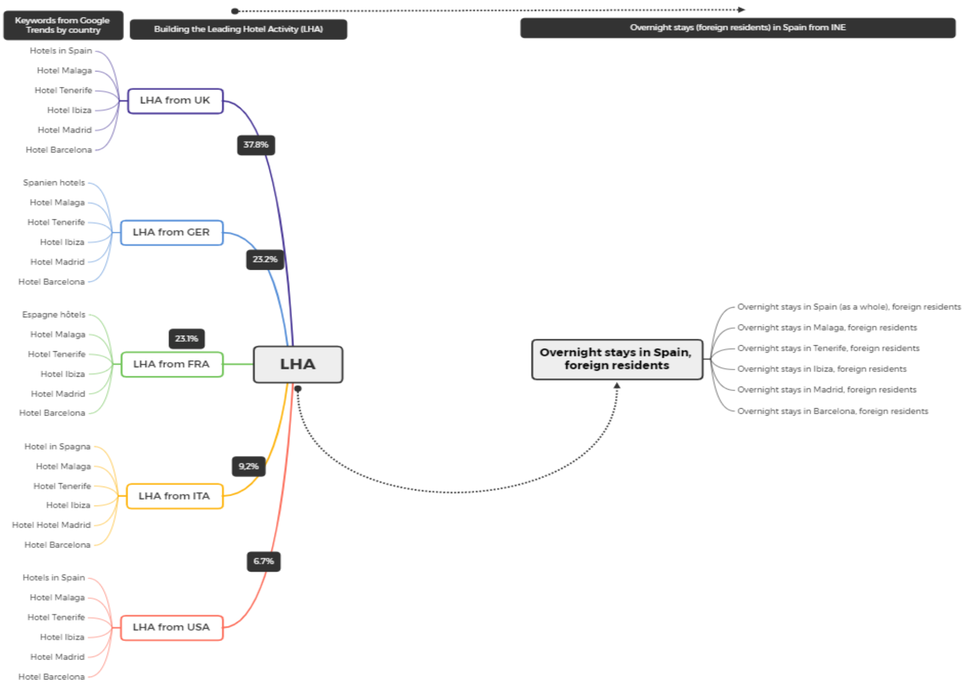
In the case of the former, we carry out a keyword search concerning the hotel sector considering the five main tourists issuing countries for Spain (the UK, Germany, France, Italy and the US) and the principal five tourist receiving regions in Spain

(Malaga, Tenerife, Ibiza, Madrid and Barcelona). As for the INE data, we analyse overnight stays of foreign residents in Spanish hotels and in the main five tourists receiving regions (Figure 1).

The Google Trends data are published on a daily basis, but we use them to build a monthly leading hotel activity index so as to compare it with the overnight stay statistics released monthly by INE. The sample period goes from January 2004 to June 2020, and therefore it includes 198 observations.

The qualitative analysis is instead based on using Padlet for online interviews of a focus group (Stewart and Shamdasani, 2017) comprising seven experts in the Spanish hotel sector, specifically two academics, one senior manager of a Destination Management Organization (DMO), a senior manager of an international travel agency and three managers from hotel companies.

Figure 1. *Concept map*



Note 1: *People from UK search keywords in Google such as hotels in Spain, hotel Malaga, hotel Tenerife, hotel Ibiza, hotel Madrid and hotel Barcelona. The weight of each keyword is as follows: 50% for Spain, 5.6% for Malaga, 20.5% for Tenerife, 8.9% for Ibiza, 10% for Madrid and 5% for Barcelona, according to the foreign visitors figures in Spain (INE, 2020). This method is used for each country (UK, Germany, France, Italy and the US).*
Note 2: *Keywords have been searched in local language (in English in UK, in French in France, and so on), but also in English for every country to reinforce the queries.*
Source: *Own elaboration based on Google Trends and INE.*

4. Empirical Results

4.1 Descriptive Analysis

The LHA index is calculated as a linear combination of the LHA indices for five countries, namely the UK, Germany, France, Italy and the US, which are based on some specific keywords (see Figure 1). The weight for each country depends on its share in total foreign visitors in Spain according to INE (2020), namely 37.8% for the UK, 23.2% for Germany, 23.1% for France, 9.2% for Italy, and 6.7% for the US.

Table 1 reports the correlations between LHA and overnight stays. It can be seen that they are positive and statistically significant, which indicates that LHA is strongly linked to hotel activity as measured by overnight stays.

Table 1. *Correlations of Overnight Stays (Foreign Residents) in Spain vs LHA*

		Overnight stays	LHA	LHA1	LHA2	LHA3	LHA6	LHA9
Overnight stays	Corr. Pearson	1	,404**	,325**	,275**	,223**	,149*	,138
	Sig. (bilateral)		,000	,000	,000	,002	,041	,060
	N	195	195	194	193	192	189	186
LHA	Corr. Pearson	,404**	1	,623**	,554**	,534**	,426**	,266**
	Sig. (bilateral)	,000		,000	,000	,000	,000	,000
	N	195	198	197	196	195	192	189
LHA1	Corr. Pearson	,325**	,623**	1	,579**	,544**	,463**	,358**
	Sig. (bilateral)	,000	,000		,000	,000	,000	,000
	N	194	197	198	197	196	193	190
LHA2	Corr. Pearson	,275**	,554**	,579**	1	,620**	,493**	,391**
	Sig. (bilateral)	,000	,000	,000		,000	,000	,000
	N	193	196	197	198	197	194	191
LHA3	Corr. Pearson	,223**	,534**	,544**	,620**	1	,534**	,426**
	Sig. (bilateral)	,002	,000	,000	,000		,000	,000
	N	192	195	196	197	198	195	192
LHA6	Corr. Pearson	,149*	,426**	,463**	,493**	,534**	1	,534**
	Sig. (bilateral)	,041	,000	,000	,000	,000		,000
	N	189	192	193	194	195	198	195
LHA9	Corr. Pearson	,138	,266**	,358**	,391**	,426**	,534**	1
	Sig. (bilateral)	,060	,000	,000	,000	,000	,000	
	N	186	189	190	191	192	195	198
** 0.01.								
* 0.05.								

Source: *Own elaboration based on Google Trends and INE.*

It can also be seen that LHA with a lead of two months exhbits a slightly stronger correlation with overnight stays, which suggests that foreigners organise their stays on average two months ahead. This evidence is consistent with that reported by Holidu (2018) and Statista (2020).

Table 2. *Correlations of Overnight Stays (Foreign Residents) in Spain vs LHA (2)*

			Overnight stays	LHA	LHA1	LHA2	LHA3	LHA6	LHA9
Spearman's Rho	Overnight stays	Correlation	1,000	,353**	,352**	,360**	,310**	,247**	,237**
		Sig. (bilateral)	.	,000	,000	,000	,000	,001	,001
		N	195	195	194	193	192	189	186
	LHA	Correlation	,353**	1,000	,694**	,631**	,611**	,471**	,332**
		Sig. (bilateral)	,000	.	,000	,000	,000	,000	,000
		N	195	198	197	196	195	192	189
	LHA1	Correlation	,352**	,694**	1,000	,651**	,633**	,512**	,399**
		Sig. (bilateral)	,000	,000	.	,000	,000	,000	,000
		N	194	197	198	197	196	193	190
	LHA2	Correlation	,360**	,631**	,651**	1,000	,688**	,555**	,439**
		Sig. (bilateral)	,000	,000	,000	.	,000	,000	,000
		N	193	196	197	198	197	194	191
	LHA3	Correlation	,310**	,611**	,633**	,688**	1,000	,611**	,471**
		Sig. (bilateral)	,000	,000	,000	,000	.	,000	,000
		N	192	195	196	197	198	195	192
	LHA6	Correlation	,247**	,471**	,512**	,555**	,611**	1,000	,611**
		Sig. (bilateral)	,001	,000	,000	,000	,000	.	,000
		N	189	192	193	194	195	198	195
	LHA9	Correlation	,237**	,332**	,399**	,439**	,471**	,611**	1,000
		Sig. (bilateral)	,001	,000	,000	,000	,000	,000	.
		N	186	189	190	191	192	195	198
** 0.01.									

Source: *Own elaboration based on Google Trends and INE.*

4.2 Quantitative Analysis

a) Univariate Analysis:

As a first step we analyse the statistical properties of the individual series, in order to establish whether or not mean reversion takes places. For this purpose, we estimate the differencing parameter d in the following model:

$$y(t) = \beta_0 + \beta_1 + x(t), \qquad (1 - L)^d x(t) = u(t), \qquad t = 1, 2, \ldots \qquad (1)$$

where y(t) is the observed time series (in our case the overnight stays and LHA series); given the monthly nature of the data, we assume that the I(0) error term in (1) follows a seasonal AR(1) process of the form:

$$u(t) = \rho\, u(t\text{-}12) + \varepsilon(t), \qquad t = 1, 2, \ldots \qquad (2)$$

where ε(t) is a white noise process.

We test the null hypothesis:

$$H_o: \; d = d_o, \qquad (3)$$

in (1) for d_o-values from 0, 0.01, … until 1.99 and 2, under three different modelling assumptions: 1) $\beta_0 = \beta_1 = 0$ a priori in (1); 2) $\beta_1 = 0$ a priori, and 3) β_0 and β_1 unknown and freely estimated from the data.

Table 3 reports the estimated values of d along with the 95% confidence intervals of the non-rejection values obtained with the test of Robinson (1994) under the assumption that u(t) is a white noise process. The estimates from the models selected on the basis of the statistical (in)significance of the deterministic terms are shown in bold. Table 4 reports the full set estimated coefficients.

Table 3. *Estimates of d Using the Series in Levels*

Series	No terms	An intercept	An intercept and a linear time trend
Overnight stays	1.02 (0.85, 1.24)	**1.25 (1.03, 1.53)**	1.25 (1.03, 1.55)
LHA	0.80 (0.70, 0.92)	0.55 (0.48, 0.66)	**0.51 (0.40, 0.69)**
LHA-UK	0.76 (0.67, 0.87)	0.45 (0.40, 0.52)	**0.35 (0.27, 0.46)**
LHA-France	0.36 (0.25, 0.48)	0.22 (0.13, 0.31)	**0.16 (0.07, 0.27)**
LHA-Germany	0.69 (0.60, 0.81)	0.43 (0.38, 0.49)	**0.31 (0.25, 0.40)**
LHA-Italy	0.70 (0.61, 0.82)	0.47 (0.41, 0.54)	**0.34 (0.25, 0.45)**
Others	0.57 (0.47, 0.69)	0.38 (0.29, 0.51)	**0.36 (0.24, 0.52)**
Stays: Malaga	0.85 (0.70, 1.08)	**0.78 (0.61, 1.03)**	0.78 (0.61, 1.03)
Stays: Balearics	0.84 (0.63, 1.12)	**1.04 (0.78, 1.31)**	1.04 (0.78, 1.31)
Stays: Madrid	0.91 (0.76, 1.09)	**0.97 (0.81, 1.15)**	0.96 (0.81, 1.15)
Stays: Barcelona	0.57 (0.43, 0.73)	0.65 (0.57, 0.77)	**0.60 (0.48, 0.75)**
Stays: Tenerife	0.99 (0.86, 1.15)	**0.70 (0.60, 0.83)**	0.69 (0.59, 0.82)

Note: *This table reports the estimates of d (with the corresoonding 95% confidence intervals in brackets) fot the three model specifications considered for the series in levels. In bold the estimates from the selected model.*
Source for LHA: *Own elaboration based on Google Trends and INE.*

Table 4. *Estimated Coefficients of the Selected Models in Table 1*

Series	d	Intercept	Time trend	Seas. AR
Overnight stays	1.25 (1.03, 1.53)	5207.24 (1.66)	---	0.963
LHA	0.51 (0.40, 0.69)	40.125 (11.52)	-0.1363 (-3.66)	0.864
LHA-UK	0.35 (0.27, 0.46)	40.184 (11.72)	-0.1224 (-4.10)	0.813
LHA-France	0.16 (0.07, 0.27)	22.874 (12.11)	-0.0558 (-3.52)	0.534
LHA-Germany	0.31 (0.25, 0.40)	35.467 (13.65)	-0.1147 (-5.18)	0.685
LHA-Italy	0.34 (0.25, 0.45)	36.661 (12.20)	-0.1608 (-6.10)	0.735
LHA-USA	0.36 (0.24, 0.52)	42.167 (11.60)	-0.1206 (-3.78)	0.515
Stays: Malaga	0.78 (0.61, 1.03)	562.839 (2.27)	---	0.980
Stays: Balearics	1.04 (0.78, 1.31)	3818.229 (1.71)	---	0.971
Stays: Madrid	0.97 (0.81, 1.15)	452.341 (1.99)	---	0.968
Stays: Barcelona	0.60 (0.48, 0.75)	488.372 (4.81)	3.216 (2.33)	0.908
Stays: Tenerife	0.70 (0.60, 0.83)	1366.925 (10.2)	--	0.891

Note: *This table reports the full set of estimates for the selected model for the series in levels, specifically those of d (with the corresponding 95% confidence intervals in brackets in column 2, and those of the intercept, the coefficient on the time trend and the seasonal coefficient (with the corresponding t-statistics in brackets for the former two) in columns 3, 4, 5, respectively).*

Source for LHA: *Own elaboration based on Google Trends and INE.*

It can be seen that the overnight stays and LHA series exhibit a different behaviour. Specifically, the time trend is significant for LHA, but not for overnight stays. Further, overnight stays are highly persistent, with a degree of integration significantly higher than 1, whilst LHA is much less persistent, and is mean-reverting ($d < 1$). The estimates of d for the LHA series are all in the interval (0, 1) and range between 0.16 (France) and 0.36 (US).

The time trend is statistically significantly negative in all cases. For overnight stays, a significant trend is found only in the case of Barcelona, and the estimated values of d range between 0.60 (Barcelona) and 1.04 (Balearic Islands); mean reversion is detected in the cases of Barcelona and Tenerife, while the unit root null cannot be rejected in the remaining three cases (Malaga, Balearic Islands, and Madrid). The last column in Table 4 shows the seasonal AR coefficient for each series – this has a large value in all cases, ranging from 0.515 (LHA-USA) to 0.980 (Stay-Malaga).

Tables 5 and 6 display the estimates of d and of the full set of coefficients respectively for the growth rate series. The time trend is now insignificant in all cases, and the estimated values of d are in the interval (0,1). For overnight stays, the estimated value of d is 0.54, and it is slightly higher (0.69) for LHA. In the latter case, the confidence interval does not include values of d smaller than 0.5, which implies that the series is non-stationary.

For the disaggregated LHA data, stationarity is found in the cases of Germany (0.29), US (0.34) and France (0.39), while for overnight stays the values are more heterogeneous, stationarity being found in the case of the Balearic Islands (0.23), and non-stationarity in the cases of Malaga (0.66) and Tenerife (0.77).

Table 5. *Estimates of d using the Growth Rate Series*

Series	No terms	An intercept	An intercept and a linear time trend
Overnight stays	**0.54 (0.42, 0.73)**	0.57 (0.44, 0.82)	0.61 (0.47, 0.85)
LHA	0.71 (0.61, 0.84)	**0.69 (0.58, 0.82)**	0.69 (0.58, 0.82)
LHA-UK	**0.48 (0.38, 0.60)**	0.48 (0.36, 0.62)	0.51 (0.39, 0.64)
LHA-France	**0.39 (0.31, 0.49)**	0.38 (0.30, 0.48)	0.38 (0.30, 0.48)
LHA-Germany	**0.29 (0.19, 0.43)**	0.28 (0.18, 0.41)	0.28 (0.18, 0.41)
LHA-Italy	0.49 (0.40, 0.60)	**0.46 (0.36, 0.57)**	0.46 (0.36, 0.57)
LHA-USA	**0.34 (0.23, 0.47)**	0.33 (0.22, 0.46)	0.33 (0.22, 0.47)
Stays: Malaga	**0.66 (0.55, 0.79)**	0.65 (0.55, 0.79)	0.66 (0.55, 0.79)
Stays: Balearics	**0.23 (0.14, 0.35)**	0.23 (0.14, 0.35)	0.23 (0.13, 0.35)
Stays: Madrid	0.55 (0.46, 0.66)	**0.56 (0.46, 0.67)**	0.56 (0.47, 0.67)
Stays: Barcelona	0.50 (0.43, 0.60)	**0.49 (0.41, 0.59)**	0.49 (0.40, 0.59)
Stays: Tenerife	**0.77 (0.67, 0.88)**	0.77 (0.68, 0.89)	0.77 (0.68, 0.89)

Note: *This table reports the estimates of d (with the corresoonding 95% confidence intervals in brackets) fot the three model specifications considered for the growth rate series. In bold the estimates from the selected model.*
Source for LHA: *Own elaboration based on Google Trends and INE.*

Table 6. *Estimated Coefficients from the Selected Models in Table 3*

Series	No terms	An intercept	An intercept and a linear time trend
Overnight stays	0.54 (0.42, 0.73)	---	---
LHA	0.69 (0.58, 0.82)	-13.138 (-1.66)	---
LHA-UK	0.48 (0.38, 0.60)	---	---

LHA-France	0.39 (0.31, 0.49)	---	---
LHA-Germany	0.29 (0.19, 0.43)	---	---
LHA-Italy	0.46 (0.36, 0.57)	-19.683 (-2.11)	---
LHA-USA	0.34 (0.23, 0.47)	---	---
Stays: Malaga	0.66 (0.55, 0.79)	---	---
Stays: Balearics	0.23 (0.14, 0.35)	---	---
Stays: Madrid	0.56 (0.46, 0.67)	6.9945 (1.67)	---
Stays: Barcelona	0.49 (0.41, 0.59)	7.7387 (1.92)	---
Stays: Tenerife	0.77 (0.67, 0.88)	---	---

Note: *This table reports the full set of estimates for the selected model for the growth rate series, specifically those of d (with the corresponding 95% confidence intervals in brackets in column 2, and those of the intercept and the coefficient on the time trend (with the corresponding t-statistics in brackets) in columns 3 and 4 respectively).*
Source for LHA: *Own elaboration based on Google Trends and INE.*

To sum up, long memory and mean reversion appear to characterise all the growth rate series; this implies that shocks have only temporary effects, which disappear at a faster rate in the case of series such as overnight stays in the Balearic Islands and LHA in Germany compared to other cases, such as overnight stays in Tenerife.

Note that the World Travel and Tourism Council managing director expressed the view in 2020 that, once the Covid-19 outbreak would be under control, it would take up to 10 months for the tourism sector to return to its normal levels (European Parliament, 2020), whilst the OECD (2020) pointed out that "eventual impacts will depend not only on the length of the pandemic, but also on potential long-term changes in travel behaviour as a result of the crisis", namely that people might be more cautious about travelling overseas in the future.

For instance, Gil-Alana and Poza (2020) found a permanent impact of the Covid-19 pandemic on the Spanish tourism sector by analysing stock market data.

b) Long-Run Equilibrium Relationships: Fractional Cointegration:

Next, we examine if there are long-run equilibrium relationships between overnight stays and LHA at an aggregated level. A necessary condition for cointegration in a bivariate setting is that the two variables should display the same degree of integration. In the case of the series in levels the orders of integration (in the first two rows in Table 4) vary from 1.25 for overnight stays to 0.51 for LHA, and the confidence intervals do not include any common values. Moreover, formal tests (Marinucci and Robinson, 2001; Hualde, 2013) reject the hypothesis of equal orders of integration.

However, in the case of the growth rate series (Table 6) the corresponding orders of integration are 0.54 and 0.69 respectively, and the null of equal orders of integration cannot be rejected at the 5% level. Therefore, next we test for cointegration between these two series.

For this purpose we use the FCVAR approach of Johansen and Nielsen (2010; 2012). The results are reported in Table 7 and imply that the hypothesis d = b cannot be rejected on the basis of a LR test, d and b being respectively the order of integration of the individual series and of the reduction in the degree of integration in the residuals from the cointegrating regression.

This implies that there exists a long-run equilibrium relationship between the two growth rate series, though the order of integration of the individual series (d = 0.932) is much higher than the one obtained through the univariate analysis. Standard cointegration tests (Johansen, 1996, with d = b = 1) produce similar results (not reported).

Table 7. *FCVAR Results*

Panel i: d ≠ b					
d	**b**	**μ_1**	**μ_2**	**1**	**2**
0.975 (0.193)	0.881 (0.221)	3.913	-5.362	0.004	0.026
Panel i: d = b					
d = b		**μ_1**	**μ_2**	**1**	**2**
0.932 (0.077)		4.201	-4.949	0.004	0.026

Notes: FCVAR *stands for Fractionally Cointegrated Vector Autoregressive model, which is used to test for long-run equilibrium relationships between fractionally integrated time series.*
Panel i: d ≠ b: This panel presents results under the assumption that the order of integration of the individual series (d) differs from b, that indicates the reduction in the degree of integration in the potential cointegration errors. In other words, the order of integration of the cointegration errors is assumed to be d – b, and in this case it is different from zero.
Panel ii: d = b: This panel imposes the restriction that the integration order in the cointegrating errors is precisely 0 as in the case of standard cointegration models.
d: Estimated order of integration of the individual time series (e.g., growth rates of LHA and overnight stays). A value between 0 and 1 indicates mean reversion and long memory.
b: Reduction in the degree of integration of the cointegration errors (residuals from the long-run relationship). Cointegration requires that b > 0.
μ_1, μ_2: Estimated intercepts in the cointegrating relationship.
1, 2: Likelihood ratio (LR) test statistics for testing model restrictions. These are used to assess the validity of the imposed constraints (e.g., d = b).
Standard errors for d and b are reported in parentheses.
Source for LHA: *Own elaboration based on Google Trends and INE.*

c) More Evidence on the Relationship between LHA and Overnight Stays:

To obtain further evidence, we assume that lagged values of LHA are weakly exogenous with respect to overnight stays, and therefore run the following regression:

$$OS(t) = \gamma_0 + \gamma_1 LHA(t-k) + x(t); \quad (1 - L)^d x(t) = u(t), \qquad t = 1, 2, \ldots \tag{4}$$

where OS(t) stands for Overnight Stays and k = 1, 2, …, 12, assuming in turn that u(t) in (4) is a white noise (Table 8) or autocorrelated (Table 9), in the latter case as in the exponential spectral model of Bloomfield (1973) which approximates ARMA structures and behaves very well in the context of fractional integration (see, e.g., Gil-Alana, 2004).

Under the assumption of white noise errors (Table 8), the estimates of d are very high for all values of k from 1 to 12, and the slope coefficient is positive and significant for k = 1, 2, 11, 12. By constrast, when allowing for autocorrelated disturbances (Table 9) the estimates of d are much lower, and even close to 0 in some cases, while the slope coefficient is now positive and significant at k = 1, 2, 3, 4, 12.

It is noteworthy that the results differ between countries (see Table A1). In particular, tourists from Germany and the UK appear to organise trips far earlier than those from the US, France and Italy. In particular, the estimates of the Germany-UK slope coefficients are positive and significant for k = 2 and 12 (Germany) and for k = 2, 4, 11 and 12 (UK).

Longer periods may be associated with leisure reasons, while shorter ones could be linked to business travel (Statista, 2020). The US, France and Italy exhibit positive and significant slope coefficients for k = 3, k = 2 and k = 1 and 2, respectively.

These results are consistent with those of Holidu (2018), who found that the tourists booking most in advance are the ones from the Netherlands, Switzerland and Germany, with an average of 77 days in advance of the arrival date, followed by those from Austria (75 days), the Anglo-Saxon countries (US, UK and Ireland – 67 days), France (60 days), Italy (52 days), Portugal (44 days), Brazil (42 days), and Spain (40 days).

Table 8. *Estimated Coefficients in the Regression of Stays on lagged LHA with white noise errors*

k	d	Intercept	Slope
1	1.68 (1.39, 1.94)	2341.658 (0.75)	**68.877 (1.87)**
2	1.71 (1.46, 1.96)	1848.521 (0.60)	**108.451 (3.00)**
3	1.78 (1.57, 1.99)	10261.683 (3.40)	-55.777 (-1.54)
4	1.75 (1.54, 1.94)	13036.110 (4.24)	-11.556 (-0.31)
5	1.76 (1.55, 1.96)	16319.318 (5.34)	-489.72 (-1.34)
6	1.77 (1.54, 1.95)	13867.070 (4.56)	59.930 (1.45)
7	1.76 (1.54, 1.99)	20136.701 (6.55)	-5.774 (-0.15)
8	1.74 (1.52, 1.97)	22341.222 (7.27)	-79.197 (-1.15)

9	1.79 (1.55, 2.01)	14474.961 (4.73)	24.780 (0.67)
10	1.85 (1.60, 2.03)	19014.291 (7.42)	-239.551 (-1.61)
11	1.78 (1.54, 1.98)	1108.996 (0.36)	**75.893 (2.11)**
12	1.68 (1.45, 1.92)	-1532.645 (-0.49)	**140.880 (3.84)**

Note: *This table reports the estimates (for different k lags and under the assumption of white noise residuals) of d (with the corresponding 95% confidence intervals in brackets) in column 2, and of the intercept and the coefficient on the time trend (with the corresponding t-statistics in brackets) in column 3 and 4 respectively. In bold the stastistically significant coefficients.*
Source for LHA: *Own elaboration based on Google Trends and INE.*

Table 9. *Estimated Coefficients in the Regression of Stays on lagged LHA with autocorrelated errors*

k	d	Intercept	Slope
1	0.38 (0.21, 0.74)	-48.601 (-0.02)	**457.044 (9.78)**
2	0.24 (0.10, 0.42)	4197.961 (3.69)	**404.044 (10.99)**
3	0.23 (0.01, 0.43)	8096.225 (6.84)	**257.736 (7.36)**
4	0.21 (-0.02, 0.78)	12630.882 (12.68)	**82.684 (2.44)**
5	0.12 (-0.09, 0.71)	17628.531 (20.29	-118.068 (-1.59)
6	0.04 (-0.14, 0.94)	21265.125 (24.97)	-273.529 (-1.04)
7	0.06 (-0.08, 0.73)	24273.449 (25.68)	-401.811 (-1.58)
8	0.06 (-0.09, 0.71)	26167.669 (28.83)	-483.816 (-1.52)
9	0.26 (0.02, 0.93)	26760.621 (17.03)	-490.911 (-1.01)
10	0.19 (0.01, 0.76)	23579.834 (24.85)	-372.798 (-1.30)
11	0.19 (-0.04, 0.97)	15494.274 (15.76)	-40.859 (-1.19)
12	0.37 (0.14, 0.91)	3718.389 (2.17)	**330.089 (7.75)**

Note: *This table reports the estimates (for different k lags and under the assumption of autocorrelated residuals) of d (with the corresponding 95% confidence intervals in brackets) in column 2, and of the intercept and the coefficient on the time trend (with the corresponding t-statistics in brackets) in column 3 and 4 respectively. In bold the stastistically significant coefficients. Source for LHA: own elaboration based on Google Trends and INE.*
Source for LHA: *Own elaboration based on Google Trends and INE.*

4.3 Qualitative Results

As already mentioned, to complement the above analysis we also use a qualitative approach based on asynchronous interviews with an online focus group (Stewart and Shamdasani, 2017).

The questions asked were the following: Q1) what will be the impact of Covid-19 on the intention to travel to Spain (temporary or more permanent) and, in particular, how will it change behaviour patterns? Q2) What measures are most effective at the company level to generate confidence in hotel stays in relation to Covid-19? Q3) What measures can be deemed as most effective at the level of authorities (national, regional, local governments) to generate confidence in hotel stays in relation to Covid-19? Q4) How long in advance, on average, do tourists from the following countries organize their trips to Spain?

Regarding the answers to Q1, there was agreement that the effects of the Covid-19 shock on the intention to travel to Spain would only be temporary; however, there was a strong consensus that patterns of behaviour would change permanently. For instance, one of the participants stated that: *"The main criteria for the selection of a tourist destination will be safety in relation to Covid-19. It might sound obvious and the challenge is how to make tangible and credible this perception of safety".*

This can be related to perceived behavioural control as one of the determinants of intentions and behaviours in the context of TPB. Moreover, most participants, mainly hotel managers, stated that they expected that tourists would perceive Spain again as a safe destination within a year. However, a hotel manager was more concerned: *"The risk comes from the comparative analysis with similar destinations such as Italy, Greece and Croatia, which could consolidate a competitive advantage. It is the better to communicate all the policies in place to protect tourist visitors."*

The DMO manager stated: *"The outcomes of the pandemic are bound to leave changes in the psychology and behaviour of the potential visitors and might even culminate into a significant change in lifestyle. In fact, 'pandemial' is a label that is beginning to be used by sociologists, anthropologists and consultants of all kinds to refer to the generation that is living the current Covid-19 pandemic, facing the complex situations that it has provoked and transforming their lives, their work, their relationships.".*

On the whole, the answers to Q1 corroborate the findings from the quantitative analysis which suggest the presence of mean reversion and thus temporary effects of the Covid-19 shock on both overnight stays and LHA (Table 6).

Regarding the answers to Q2, participants shared a wide range of practices that were adopted by all the main hotels (masks, hand sanitizers, capacity restrictions in commons spaces such as restaurants, swimming-pools, cash free, temperature check-in devices, etc.). It was interesting to find that employees' compliance with rules and procedures was the highest challenge.

For instance, a hotel manager said: *"My main energies are concentrated in our employees' behaviour in relation to Covid-19 safeguarding measures. If they comply and are committed, this will have a positive impact on customers because this is a matter of exemplarity and persuasion. But mainly our energies are there because we want first and foremost to protect our employees because we really care for them."*

One academic who had done extensive research on tourism and was regularly in contact with industry managers stated: *"It is important to understand that safety* *works as a hygiene factor, meaning that when it is present it does not increase the perception of quality, but when it is absent it produces a perception of bad quality. Therefore, it is absolutely essential to have the necessary safety measures for guests to make them feel secure.".*

As for the answers to Q3, there was no consensus about the importance of the measures adopted by the authorities. Some interviewees argued that what mattered most was that hotels should communicate clearly what measures they had in place, though only for beach resorts where clients were planning to stay mainly on the hotel premises and not visit the surroundings.

Conversely, others took the view that the role of authorities was critical to generate confidence in a destination. For instance, one interviewee said: *"It is important that authorities invest a huge amount of money in promotional campaigns in order to reinforce the image of a safe destination."* Another participant stressed the importance of economic and financial measures including *"soft loans, tax relief such as a clear reduction in VAT for tourism, and specific and generous furlough schemes for hotel operators.".*

A hotel manager expressed some concern: "*Until there is a vaccine, there needs to be tests at the origin and destination, but outside of this, there needs to a serious educational communication campaign around the taking of this vaccine, there is a lot of fake news and negative comments around the taking of this vaccine by senior politicians, which will impact people's willingness to take it and this will impact the return of our economies.".*

The DMO manager was very clear about his opinion regarding the authorities: *"In my view they have made many mistakes including two ajor ones: after the first wave, they relaxed restrictions very quickly and this gave a wrong message that the nightmare was over and party time was back and the other mistake was to be permissive with the nightlife sector".*

A hotel director said: *"The problem is that many destinations, such as Spain, tried to transmit security in their marketing messages (secure corridors), but they did not couple those statements with appropriate measures."*

Another participant responded to this comment as follows: *"In the final analysis, that is one of the reasons why the second wave of the pandemic has started earlier and more acutely in some destinations than in others. We can, for example, compare the case of Italy, which demanded security measures to incoming travelers and had important restrictions for bars and restaurants, such as PCRs, or Spain, that did not. Even in the case of the latter, when tougher measures, such as the mandatory use of masks, were imposed to curb the rapid increase in contagion, the hospitality sector openly opposed them. The result is that Spain went on to lead the coronavirus statistics in Europe, and that the tourism season came abruptly to an end at the* *beginning of September, when the majority of countries imposed restrictions to travel to Spain.”.*

Most of participants agreed that health should be put before the economy. The following comment reflects well a general consensus: *“One of the lessons of this pandemic is that you cannot put the economy before health. When this is done, the health problems it causes have even worse economic consequences.”.* One participant responded with this comment*: “compulsory PCRs for anybody coming into the destination and regular PCRs for all the employees are, in my opinion, a must. And very strict measures on bars and restaurants (not the ones we had in Spain) with a lot of control, inspection and heavy fines.”.*

Concerning the answers to Q4, there was a general consensus that hotel visitors from Germany and the UK were those who planned most in advance their visits to Spain, followed by the US, France and Italy. On average, visitors from Germany and the UK arranged their visits six or more months ahead, those from the US six months ahead, and those from France and Italy less than six months ahead. Participants discussed the importance of segmentation for reaching valid conclusions.

Additional comments shed more light on this issue. For instance, one participant stated: *“If we consider business travelers, normally they book their hotel rooms with short notice and this is mainly managed by their own companies and I expect that they will be the less affected by Covid-19. If you need to travel for business, you don´t cancel it easily. Another important segment is families that come during half term school vacations. These are the most predictable, in particular Britons, who plan a lot of time in advance as they know the dates of half term before the academic year starts. This segment will be more affected by Covid-19 and staycation in the UK will be our main competitor.”.*

Again this evidence from the qualitative study is consistent with our previous quantitative findings (and also with those of Holidu, 2018), as the regressions analysing the relationship between LHA and overnight stays also yield positive and significant slope coefficients at longer lags for countries such as Germany, the UK etc.

5. Conclusions

The main aim of this study was to assess the impact of the Covid-19 shock on Spanish hotels activity and to establish whether the effect would be temporary or permanent. For these purposes we used both quantitative methods (time series analysis) and qualitative ones (online interviews).

We also constructed a leading hotel activity index (LHA) to examine its relationship with foreigners’ overnight stays and assess its usefulness to predict future trends in the latter, and carried out the analysis separately for the five main tourist issuing countries for Spain (the UK, Germany, France, Italy and the US) and the principal five tourist receiving regions in Spain (Malaga, Tenerife, Ibiza, Madrid and Barcelona).

In brief, our findings indicate the following:

- The Leading Hotel Activity (LHA) index, based on Google Trends, exhibits a positive and statistically significant correlation with foreign resident overnight stays in Spain. This correlation increases with a lead of two months in LHA, probably because of the time required to organise the stays.
- The growth rates of LHA and overnight stays are cointegrated and thus the former contains useful information about the latter.
- The effects of the Covid-19 shock on Spanish hotel activity were significant and negative but only temporary, and disappeared at a faster rate in the case of series such as overnight stays in the Balearic Islands and of tourists from Germany compared to other cases, such as overnight stays in Tenerife and tourists from the UK.
- Foreign residents organise their stays one and two months ahead in some cases and eleven and twelve months in others. One possible explanation for these differences is the type of stay: if it is for leisure purposes, it tends to be organised well in advance, whilst if it is related to business it is often planned at short notice due to demand inelasticity (Statista, 2020). Also, there are behavioural differences between countries; specifically, German and British visitors usually organise their stay earlier than American, French and Italian visitors. However, these factors cannot be tested formally using our data, which is a limitation of our research.

On the whole, this study provides new evidence on the relationships between intentions to visit a destination and stay in a hotel and the subsequent final visit. Both the quantitative and the qualitative findings confirm that intentions can be a predictor of behaviour, as suggested by TPB.

In particular, they reinforce the conclusions of the study by Moon (2021) where perceived behavioural control is the main determinant of intentions and behaviour. Other factors, such as creating “safe havens” in the hospitality industry (Hu et al., 2020) and the correct implementation of internal market orientation (Ruizalba, et al. 2014), can also have an impact.

Future research could examine additional source markets for Spain, since the five we have considered (the UK, Germany, US, France and Italy) only cover 60% of Spanish tourism (INE, 2020).

Also, although Spain is the second tourist destination in the world (UNWTO, 2020), extending the analysis to other destination countries would yield further insights into the issues of interest.

References:

Ajzen, I. 1991. The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), pp. 179-211. https://doi.org/10.1016/0749-5978(91)90020-T.

Antenucci, D.,Cafarella, M., Levenstein, M., Ré, Che and Shapiro, M.D. 2014. Using Social Media to Measure Labor Market Flows. Working Paper No. 20010. National Bureau of Economic Research, March.

Anton, C.E., Lawrence, C. 2016. The relationship between place attachment, the theory of planned behaviour and residents' response to place change. Journal of Environmental Psychology. http://dx.doi.org/10.1016/j.jenvp.2016.05.010.

Bangwayo-Skeete, P.F., Skeete, R.W. 2015. Can Google data improve the forecasting performance of tourist arrivals? Mixed-data sampling approach. Tourism Management, 46, pp. 454-464. https://doi.org/10.1016/j.tourman.2014.07.014.

Bernanke, B.S. 2008. Stabilizing the financial markets and the economy. Speech at the Economic Club of New York.

Bettman, J.R. 1979. An Information Processing Theory of Consumer Choice. Reading, MA.: Addison-Wesley.

Bianchi, C., Milberg, S., Cúneo, A. 2017. Understanding travelers' intentions to visit a short versus long-haul emerging vacation destination: The case of Chile. Tourism Management, 59, 312-324. https://doi.org/10.1016/j.tourman.2016.08.013.

Bloomfield, P. 1973. An exponential model in the spectrum of a scalar time series. Biometrika, 60, 217-226. https://doi.org/10.1093/biomet/60.2.217.

Bollen, J., Mao, H., Zeng, X. 2011. Twitter mood predicts the stock market. Journal of Computational Science, 2(1), pp. 1-8. https://doi.org/10.1016/j.jocs.2010.12.007.

Caporale, G.M., Gil-Alana, L.A., Poza, C., Ruiz-Alba, J.L. 2025. The COVID-19 Shock and Spanish Hotel Activity (No. 11985). CESifo Working Paper.

Choi, H., Varian, H. 2012. Predicting the present with Google Trends. Economic Record, 88(s1), pp. 2-9. https://doi.org/10.1111/j.1475-4932.2012.00809.x.

Chung, N., Lee, H., Lee, S.J., Koo, C. 2015. The influence of tourism website on tourist's behavior to determine destination selection. Technology Forecasting and Social Change, 96, pp. 130-143. https://doi.org/10.1016/j.techfore.2015.03.004.

Creswell, J.W. 2013. Research Design: Qualitative, Quantitative, and Mixed Methods approaches. Sage.

Cutler, S.Q., Carmichael, B.A. 2010. The dimensions of the tourist experience. The tourism and leisure experience: Consumer and Managerial Perspectives, 44, 3-26.

D'Amuri, F., Marcucci, J. 2017. The Predictive Power of Google Searches in forecasting US unemployment. International Journal of Forecasting, 33(4), pp. 801-816. https://doi.org/10.1016/j.ijforecast.2017.03.004.

Dong, L., Chen, S., Cheng, Y., Wu, Z., Li, C., Wu, H. 2017. Measuring economic activity in China with mobile big data. EPJ Data Science, 6(29), pp. 1-17.

Erawan, T., Krairit, D., Khang, D.B. 2011. Tourists' external information search behavior model: The case of Thailand. Journal of Modelling in Management, 6(3), 297-316. DOI: 10.1108/17465661111183701.

European Parliament. 2020. COVID-19 and the tourism sector. European Parliamentary Research Service.

Gil-Alana, L.A. 2004. The use of Bloomfield (1973) model as an approximation to ARMA processes in the context of fractional integration. Mathematical and Computer Modelling 39, 429-436. https://doi.org/10.1016/S0895-7177(04)90515-8.

Gil-Alana, L.A., Poza, C. 2020. The impact of COVID-19 on Spanish Tourism Sector. Tourism Economics 28, Vol. 3. https://doi.org/10.1177/13548166209599.
Han, H., Hsu, L.T., Sheu, C. 2010. Application of the Theory of Planned Behavior to green hotel choice: Testing the effect of environmental friendly activities. Tourism Management, 31(3), 325-334. https://doi.org/10.1016/j.tourman.2009.03.013.
Havranek, T., Zeynalov, A. 2021. Forecasting tourist arrivals: Google Trends meets mixed-frequency data. Tourism Economics, 27(1), 129-148. https://doi.org/10.1177/1354816619879584.
Hisano, R., Sornette, D., Mizuno, T., Ohnishi, T. 2013. High quality topic extraction from business news explains abnormal financial market volatility. PLoS ONE, 8(6), e64846. https://doi.org/10.1371/journal.pone.0064846.
Hu, X., Yan, H., Casey, T., Wu, C.H. 2020. Creating a safe haven during the crisis: How organizations can achieve deep compliance with COVID-19 safety measures in the hospitality industry. International Journal of Hospitality Management, 102662. https://doi.org/10.1016/j.ijhm.2020.102662.
Hualde, J. 2013. A simple test for the equality of integration orders. Economics Letters, 119(3), pp. 233-237. https://doi.org/10.1016/j.econlet.2013.03.003.
Huang, M.Y., Rojas, R.R., Convery, P.D. 2018. News Sentiment as Leading Indicators for Recessions. Department of Economics. University of California, USA.
Johansen, S. 1996. Likelihood-based inference in cointegrated vector autoregressive models. New York, NY, Oxford University Press.
Johansen, S., Nielsen, M.Ø. 2010. Likelihood inference for a nonstationary fractional autoregressive model. Journal of Econometrics, 158(1), pp. 51-66. https://doi.org/10.1016/j.jeconom.2010.03.006.
Johansen, S., Nielsen, M.Ø. 2012. Likelihood inference for a fractionally cointegrated vector autoregressive model. Econometrica, 80(6), pp. 2667-2732. https://doi.org/10.3982/ECTA9299.
Kaushal, V., Srivastava, S. 2021. Hospitality and Tourism Industry amid COVID-19 Pandemic: Perspectives on Challenges and Learnings from India. International Journal of Hospitality Management, 102707. https://doi.org/10.1016/j.ijhm.2020.102707.
Kim, Y.J., Njite, D., Hancer, M. 2013. Anticipated emotion in consumers' intentions to select eco-friendly restaurants: Augmenting the theory of planned behavior. International Journal of Hospitality Management, 34, pp. 255-262. DOI: 10.1016/j.ijhm.2013.04.004.
Kim, E., Lee, J.H., Sung, Y., Choi, S.M. 2016. Predicting selfie-posting behavior on social networking sites: An extension of theory of planned behavior. Computers in Human Behavior, 62, pp. 116-123. DOI: 10.1016/j.chb.2016.03.078.
Koenig, E. 2002. Using the purchasing managers' index to assess the economy's strength and the likely direction of monetary policy. Economic and Financial Policy Review, Federal Reserve Bank of Dallas, 1(6), pp. 1-14.
Li, X., Pan, B., Law, R., Huang, X. 2017. Forecasting tourism demand with composite search index. Tourism Management, 59, pp. 57-66. https://doi.org/10.1016/j.tourman.2016.07.005.
Marinucci, D., Robinson, P.M. 2001. Semiparametric fractional cointegration analysis. Journal of Econometrics, 105(1), pp. 225-247. https://doi.org/10.1016/S0304-4076(01)00076-8
Moon, S.J. 2021. Investigating beliefs, attitudes, and intentions regarding green restaurant patronage: An application of the extended theory of planned behavior with

moderating effects of gender and age. International Journal of Hospitality Management, Volume 92, 102727. DOI: 10.1016/j.ijhm.2020.102727.

Nyman, R., Ormerod, P., Smith, R., Tuckett, D. 2014. Big Data and Economic Forecasting: A Top-Down Approach Using Directed Algorithmic Text Analysis. Workshop on Big Data for Forecasting and statistics. ECB, Frankfurt.

OECD. 2020. Tourism Policy Responses to the coronavirus (COVID-19). OECD Policy Responses to Coronavirus (COVID-19).

Önder, I., Gunter, U. 2016. Forecasting Tourism Demand with Google Trends For a Major European City Destination. Tourism Analysis, 21, 203-220. DOI: 10.3727/108354216X14559233984773.

Padhi, S.S., Pati, R.K. 2017. Quantifying potential tourist behavior in choice of destination using Google Trends. Tourism Management Perspectives, 24, pp. 34-47. DOI: 10.1016/j.tmp.2017.07.001.

Pappalardo, L., Vanhoof, M., Gabrielli, L., Smoreda, Z., Pedreschi, D., Giannotti, F. 2016. An anaytical framework to nowcast well-being using mobile phone data. International Journal of Data and Scientific Analysis, 2 (1-2), pp. 75-92. https://doi.org/10.1007/s41060-016-0013-2.

Poza, C., Monge, M. 2020. A real time leading economic indicator based on text mining for the Spanish economy. Fractional cointegration VAR and Continuous Wavelet Transform analysis. International Economics, 163, pp. 163-175. DOI: 10.1016/j.inteco.2020.02.002.

Quintal, V.A., Lee, J.A., Soutar, G.N. 2010. Risk, uncertainty and the theory of planned behavior: A tourism example. Tourism Management, 31, 797-805. https://doi.org/10.1016/j.tourman.2009.08.006.

Rivera, R. 2016. A dynamic linear model to forecast hotel registrations in Puerto Rico using Google Trends data. Tourism Management, 57, pp. 12-20. https://doi.org/10.1016/j.tourman.2016.04.008.

Rivera, M. 2020. Hitting the reset button for hospitality research in times of crisis: Covid-19 and beyond. International Journal of Hospitality Management. Vol, 87. DOI: 10.1016/j.ijhm.2020.102528.

Robinson, P.M. 1994. Efficient tests of nonstationary hypotheses. Journal of the American Statistical Association. 89, pp. 1420-1437. https://doi.org/10.1080/01621459.1994.10476881.

Ruizalba, J.L., Bermúdez-González, G., Rodríguez-Molina, M.A., Blanca, M.J. 2014. Internal market orientation: An empirical research in hotel sector. International Journal of Hospitality Management, 38, 11-19. DOI: 10.1016/j.ijhm.2013.12.002.

Scott, S.L., Varian, H.R. 2012. Bayesian Variable Selection for Nowcasting Economic Time Series. December 31 Draft at Sebastiani, F. 2002. Machine learning in automated text categorization. ACM Computing Surveys (CSUR) Surveys, 34(1), pp. 1-47.

Stewart, D.W., Shamdasani, P. 2017. Online focus groups. Journal of Advertising, 46(1), pp. 48-60. https://doi.org/10.1080/00913367.2016.1252288.

Sun, Sh., Wei, Y., Tsui, K.L., Wang, Sh. 2019. Forecasting tourist arrivals with machine learning and internet search index. Tourism Management, 70, pp. 1-10. DOI: 10.1016/j.tourman.2018.07.010.

Szármes, P. 2015. Creating new types of business and economic indicators using big data technologies. Science Journal of Business and Management. 3(1-1), pp. 18-24.

Toole, J.L., Lin, Y.R., Muehlegger, E., Shoag, D., Gonzalez, M.C., Lazer, D. 2015. Tracking employment shocks using mobile phone data. Journal of The Royal Society Interface, 12 (107).

Varian, H.R. 2014. Big Data: new tricks for econometrics. Journal of Economic Perspectives, 28(2), pp. 3-27. DOI: 10.1257/jep.28.2.3.

Yang, X., Pan, B., Evans, J.A., Lv, B. 2015. Forecasting Chinese tourist volume with search engine data. Tourism Management, 46, pp. 386-397. https://doi.org/10.1016/j.tourman.2014.07.019.

Internet Sources:

Holidu. 2018. Tu buscador de alquileres vacacionales. https://www.holidu.es/.

INE. 2020. Datos por temas. https://www.ine.es/.

Statista. 2020. How far in advance will or have you book(ed) your main holiday? https://www.statista.com/statistics/320998/lead-time-between-holidy-booking-and-departure-uk/.

UNWTO. 2020. UNWTO Tourism Data Dashboard. World Tourism Organization. https://www.unwto.org/es/unwto-tourism-dashboard.

Table A1. *Estimated coefficients in the regression of stays on lagged LHA: Issuing countries (France – Italy – Germany – UK – USA)*

	FRANCE			ITALY			GERMANY			UK			USA		
k	d	Intercept	Slope	d	Intercept	Slope	d	Intercept	Slope	d	Intercept	Slope	d	Intercept	Slope
1	1.74 (1.51, 1.97)	4268.91 (1.43)	19.9852 (0.92)	1.69 (1.43, 1.93)	3134.18 (1.18)	**77.4920 (2.76)**	1.72 (1.47, 1.97)	5100.01 (1.96)	23.4335 (0.87)	1.71 (1.45, 1.96)	4430.66 (1.63)	27.9751 (1.23)	1.75 (1.53, 1.98)	5967.45 (2.29)	-2.6234 (-0.12)
2	1.73 (1.52, 1.96)	4025.36 (1.35)	**42.6160 (1.97)**	1.70 (1.45, 1.94)	5070.68 (1.91)	**65.9200 (2.34)**	1.74 (1.50, 1.97)	5205.03 (2.04)	**68.5473 (2.61)**	1.74 (1.47, 1.98)	5198.21 (1.94)	**43.2176 (1.93)**	1.77 (1.54, 1.99)	8361.49 (13.24)	-21.3271 (-1.01)
3	1.75 (1.53, 1.97)	7865.89 (2.64)	-5.4090 (-0.25)	1.74 (1.52, 1.97)	6548.66 (2.48)	25.6834 (0.91)	1.76 (1.53, 1.98)	8200.68 (3.19)	-24.3303 (-0.91)	1.79 (1.51, 2.00)	10263.24 (3.96)	-56.6448 (-1.60)	1.72 (1.50, 1.94)	4939.88 (1.89)	**55.9991 (2.63)**
4	1.76 (1.55, 1.98)	15757.14 (5.36)	-41.7540 (-1.85)	1.78 (1.57, 2.00)	15345.17 (6.07)	-83.2338 (-1.09)	1.75 (1.54, 1.97)	13669.63 (5.32)	-38.4413 (-1.44)	1.78 (1.54, 2.02)	10278.04 (3.90)	**42.5304 (1.91)**	1.72 (1.51, 1.94)	11269.43 (42.51)	26.2684 (1.21)
5	1.75 (1.52, 1.98)	14829.49 (4.97)	-12.6781 (-0.58)	1.75 (1.53, 1.98)	13488.17 (5.12)	9.6242 (0.34)	1.74 (1.54, 1.98)	14285.44 (5.52)	-14.4615 (-0.53)	1.78 (1.54, 2.01)	15924. 32 (6.03)	-41.9348 (-1.08)	1.74 (1.53, 1.97)	12571.63 (4.78)	28.0005 (1.30)
6	1.76 (1.55, 1.99)	14586.37 (4.92)	29.6967 (1.37)	1.77 (1.56, 1.99)	15521.60 (5.97)	40.3622 (1.45)	1.76 (1.55, 1.99)	15903.51 (6.19)	32.7995 (1.23)	1.76 (1.53, 1.98)	15928.62 (5.92)	20.0055 (0.88)	1.74 (1.52, 1.97)	17120.22 (6.48)	-33.9922 (-0.15)
7	1.76 (1.54, 1.99)	20926.30 (7.02)	-1366.54 (-0.62)	1.75 (1.52, 1.98)	21135.35 (8.05)	-36.9391 (-1.32)	1.76 (1.55, 1.97)	19463.30 (7.55)	11.8935 (0.44)	1.76 (1.53, 1.98)	19238.57 (7.15)	11.8773 (0.52)	1.76 (1.54, 1.99)	21642.14 (8.35)	-39.9577 (-1.09)
8	1.76 (1.54, 1.98)	20312.90 (6.81)	-25.3460 (-1.17)	1.73 (1.50, 1.98)	19878.41 (7.50)	-4533.83 (-1.60)	1.76 (1.54, 1.98)	19258.73 (7.47)	-30.1366 (-1.12)	1.76 (1.53, 1.98)	20605.41 (7.71)	-45.4095 (-1.01)	1.79 (1.55, 2.01)	17264.53 (6.69)	23.1894 (1.10)
9	1.76 (1.53, 1.99)	15899.73 (5.29)	-2.7528 (-0.12)	1.73 (1.50, 1.97)	17354.28 (6.53)	-49.6367 (-1.15)	1.76 (1.52, 2.00)	15750.12 (6.06)	-2.1686 (-0.08)	1.87 (1.56, 2.19)	13339.38 (5.23)	50.4435 (2.35)	1.73 (1.52, 1.97)	17850.57 (6.55)	-49.5000 (-1.29)
10	1.75 (1.54, 1.98)	7490.18 (2.51)	-10.2821 (-0.47)	1.75 (1.51, 1.98)	7118.74 (2.71)	-12.6711 (-0.45)	1.78 (1.53, 2.00)	10656.06 (4.46)	-12.51138 (-1.05)	1.92 (1.74, 2.14)	15195.56 (7.43)	-16d6.24 (-1.46)	1.73 (1.52, 1.97)	9040.11 (3.47)	-52.6504 (-1.47)
11	1.75 (1.53, 1.98)	3756.85 (1.25)	15.5607 (0.71)	1.75 (1.50, 1.98)	4991.63 (1.89)	-0.0504 (-0.01)	1.76 (1.54, 2.00)	3948.11 (1.54)	32.7964 (1.23)	1.79 (1.56, 2.02)	2422.23 (0.92)	**50.6639 (2.30)**	1.76 (1.53, 1.98)	4748.34 (1.81)	5.3785 (0.25)
12	1.75 (1.52, 1.98)	4970.35 (1.65)	8.7851 (0.40)	1.74 (1.49, 1.97)	4968.05 (1.87)	19,8858 (0.70)	1.72 (1.50, 1.94)	3405.92 (1.31)	**70.3693 (2.62)**	1.70 (1.47, 1.92)	1024.42 (0.38)	**91.1953 (4.07)**	1.76 (1.52, 1.98)	4177.03 (1.60)	33.1128 (1.55)

Note: *This table reports for each country the estimates (for different k lags and under the assumption of autocorrelated residuals) of d (with the corresponding 95% confidence intervals in brackets), and of the intercept and the coefficient on the time trend (with the corresponding t-statistics in brackets). In bold the stastistically significant coefficients. Source for LHA: own elaboration based on Google Trends and INE.*

Source for LHA: *own elaboration based on Google Trends and INE.*