ON THE ANALYSIS OF EFFICIENCY IN THE HOTEL SECTOR:

does tourism specialization matter?

Authors:

Federico Inchausti-Sintes[1]

Ubay Pérez-Granja[2]

Accepted for publication in Tourism Economics:
https://doi.org/10.1177/13548166211039301

Abstract

This paper analyzes the consequences of tourism specialization on efficiency in the hotel sector. The evidence found in other sectors and economies supports the goodness specialization. Nevertheless, tourism-led economies have particular issues that need to be addressed such as: seasonality and the lack of significant tradable competitive activities that could trigger spillover effects to services. Spain provides a suitable context for a comparative case study where industrial-led provinces coexist with others that are tourism-led. The paper assumes a novel panel-data Stochastic Frontier model where inefficiency is explained by industrial and service specialization, international competitiveness, tourism specialization, quality of tourism supply, and seasonality. All variables contribute to reducing inefficiency, but service specialization makes the biggest impact. Hence, tourism-led provinces produce the highest efficiency scores.

Keywords: Stochastic frontier, efficiency, service specialization, industrial specialization and seasonality

[1] Federico Inchausti-Sintes, Facultad de Economía, Empresa y Turismo, Universidad of Las Palmas de Gran Canaria, C/ SauloTorón s/n, Despacho D. 2.15, CP 35017 Las Palmas de Gran Canaria, Spain. Email: federico.inchausti@ulpgc.es

[2] Ubay Pérez-Granja, Facultad de Economía, Empresa y Turismo, Universidad of Las Palmas de Gran Canaria, C/ SauloTorón s/n, Despacho D. 2.15, CP 35017 Las Palmas de Gran Canaria, Spain. Email: ubay.perez@ulpgc.es

1. Introduction

As comprehensively explained by Torrens (1815) and Ricardo (1817), two economies can obtain a mutual benefit by specializing in a good with a comparative advantage and trading to the other. As a result, both economies are capable of consuming above their respective production possibility frontiers. Likewise, Smith (1789) also noted the goodness of specialization in productivity through the process of dividing one task into successively smaller ones, where each worker specializes, instead of trying to accomplish the entire task on their own. In both cases, specialization makes more efficient use of resources and increases competitiveness. Apart from technological reasons, as detailed by Ricardo (1817), the comparative advantage can also be triggered by the physical proximity between economies or the availability of resources in the territory (Feenstra & Taylor, 2011). Bougheas, Demetriades and Manuneas (2000) also demonstrate how infrastructures can facilitate economic specialization when they are employed as a cost-reducing technology. In the long term, specialization also leads to spatial agglomeration by concentrating those activities with the same comparative advantages (Krugman & Venables, 1996), and it enhances economic growth, especially when focusing on goods with high-technological content (Lee, 2011). In the case of tourism specialization, the availability of natural resources (climate and beaches, mainly) plays a determinant role in this process, while, at the same time, boosting the spatial agglomeration of tourism activities nearby (Eugenio-Martín, Cazorla-Artiles & González-Martel, 2019). Likewise, tourism specialization also leads to economic growth (Brau, Lanza & Pigiaru, 2007) and seems to be more intense in countries with a high economic level and financial development (De vita & Kyaw, 2017).

Based on this evidence, one might presume that tourism-led economies provide more efficient use of their resources in the hospitality sector; precisely because of their service-based orientation. However, specialization in tourism-led economies poses other specific issues that need to be addressed. Firstly, hospitality service demand commonly experiences sharper fluctuations, which intensify at tourism destinations because of the seasonality of tourism flows. Firms tend to hire part-time workers to adapt their capacity in such circumstances, which reduces the opportunity for workers to acquire workplace skills, in addition to the fact that these firms are more reluctant to invest in training. As a result, both the level and quality of the services delivered weaken; eventually leading to sub-optimal long-term profits (Alemayehu & Tveteraas, 2019). Within the hospitality sector, hoteliers might be more vulnerable to fluctuations in demand; because they have to deal with a greater magnitude off fixed capital to be matched with both labour and demand. Secondly, both industrial and service activities benefit from outsourcing some of their processes to allow for competitivity gains, as noted by Fixler & Siegel (1999). However, tourism predominantly relies on a specific kind of service (non-tradable) that cannot easily be outsourced. Finally, tourism-led economies have a productive-mix that is highly focused on services and with a marginal share of industrial activities (Parrilla, Font & Nadal, 2007; and Inchausti-Sintes, 2019). This kind of productive-mix limits, for instance, the linkage and potential spillover effects from more technological industries to more labour-intensive services. In some cases, the aforementioned economic consequences may be considered as a symptom of "dutch disease" in tourism-led economies (Chao, Hazari, Laffargue, Sgro, & Eden, 2006; Nowak, & Sahli, 2007; or Capo, Font, & Nadal, 2007). In sum, these kinds of issues may refute our initial supposition regarding the goodness of service specialization in tourism-based economies.

Specifically, this paper analyzes whether specialization in the hotel sector in tourism-led economies leads to greater efficiency than in other kinds of economies. The analysis focuses on the 50 Spanish provinces (NUTS III) during the period 2001-2016. These provinces provide a suitable comparative case study because more technologically-led and economically diversified provinces (such as Álava, Guipuzcoa, Vizcaya or Navarre), coexist with more traditional provinces (Seville, Badajoz, León or Albacete) and well-known tourism-led economies such as the Balearic Islands, the Canary Islands and some Mediterranean provinces.

Methodologically, the paper uses an annual panel data Stochastic Frontier (SF) model developed by Lien, Kumbhakar and Alem (2018). This model extends previous works (Kumbhakar, Lien & Hardaker, 2014; and Colombi, 2014) by allowing determinants in the inefficiency term and dealing with endogeneity, which has gone unaddressed in most SF analysis. In this sense, 'efficiency' is modeled by controlling for both service and other sectoral specialization that may trigger a favourable spillover effect; as well as seasonality. Finally, the model also distinguishes three components: heterogeneity; permanent and transient inefficiency components.

In sum, all provinces report high-efficiency scores, but service specialization plays a key role in reducing inefficiency in the hotel sector; with the two leading tourism-led economies reporting the highest efficiency scores. Industrial specialization also has a positive impact on hotel efficiency by easing linkage and spillover effects, but much lower than the impact of the aforementioned specialization. The study produced an additional novel result by showing that, on average, lower seasonality does not necessarily mean more efficiency.

The remainder of this paper is structured as follows: Section 2 reviews the literature on the analysis of efficiency in tourism. Section 3 covers the case-study and section 4 explains the methodology. Section 5 presents the results and discusses their main implications. Section 6 concludes with a summary of the main findings.

2. Literature review

Efficiency analysis in tourism has increased in popularity over the past two decades. Two methodologies have monopolized the efficiency literature in tourism: Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis (SFA). However, the former has overtaken the latter in terms of use. According to Assaf and Josiassen's meta-analysis (2016), out of 57 studies analyzed, 49 were micro studies and only eight were at macro level. Moreover, the majority of the tourism literature has focused predominantly on hotels. According to the aforementioned analysis, 43 applied different forms of DEA, with only 14 papers employing SFA. In fact, in recent years the trend seems to be the same, with the DEA methodology being more used than SFA.

The rest of this section is organized as follows. The first two subsections focus on analyzing the application of DEA and SFA at micro level (*micro-level studies*) and macro level (*macro-level studies*). The following subsection concentrates on studying those papers that, using either of these two methodologies, take Spain as their case-study.

Subsequently, the next subsection analyzes the literature on SFA. Finally, Table 1 summarizes all the applied papers mentioned during the whole section.

Micro-level studies

In recent years, the number of micro-level studies on efficiency in tourism has fallen. However, the topic and has been widely applied in different tourism activities. For instance, Ben Aissa and Goaied (2016) use the managerial efficiency scores obtained through a DEA to analyze their impact on hotel profitability, and found a positive relationship. Chang, Lee and Park (2017), on the other hand, measured the efficiency of the major cruise lines using a DEA. The authors found that the lines attempting high capacity expansion were more inefficient in relative terms. Deng, Gu, Law and Lian (2020) measured the efficiency of casinos in Macao and Las Vegas and found that there were completely different efficiency drivers at each destination. The relationship between sustainability and efficiency has been also studied: Kularatne, Wilson, Mansson, Hoang and Lee (2019), for example, used a DEA to assess efficiency and its determinants, and concluded that most environmentally responsible actions led to improvements in efficiency. The role of tourism in airports' technical efficiency has also been studied by Ripoll-Zarraga and Raya (2020) and Fernández, Coto-Millán and Díaz-Medina (2018) using SFA. Both papers found that tourism-oriented airports are more efficient than non-touristic ones.

Macro-level studies

At macro level, the literature has been more extense in comparison to the situation described in Assaf and Josiassen (2016) (see for example: Cuccia, Guccio & Rizzo, 2017, Karakitsiou, Kourgiantakis, Mavrommati & Migdalas, 2018; or Zhou, Xu & Lee, 2019 among others). Specifically, there have been studies which focused on the determinants of efficiency. Cuccia, Guccio and Rizzo (2017), for example, applied a DEA that aimed to study the effect of being on the World Heritage List (UNESCO) on the efficiency performance of Italian tourism destinations. However, the authors could not identify any advantage for those on the list nor significant spillover effects in the efficiency of Italian tourism destinations. Huang (2018) employed a hybrid network DEA to evaluate the performance of 30 tourism supply chains in China. Song and Li (2019) used a DEA to estimate the efficiency of the Chinese tourism industry. The authors analyzed the determinants of efficiency and concluded that economic development, urbanization and openness have a positive impact on efficiency. Nevertheless, pure efficiency analysis is still a relevant issue in the literature, for example Karakitsiou et al. (2018) used a DEA model to estimate the efficiency of the hospitality sector (hotels and restaurants).

Spain as a case-study.

Spain has been prolifically employed as a case study in the efficiency literature (see for example: Alberca-Oliver, Rodríguez-Oromendia & Parte-Esteban, 2015; Arbelo-Pérez, Pérez-Gómez & Arbelo, 2019; Benito, Solana & López, 2014; Pérez-Rodríguez & Acosta-González, 2007; Sellers-Rubio & Casado-Díaz, 2018; Solana-Ibañez, Caravaca-Garratón & Para-González, (2016a) and (2016b) among others).

At micro-level, the literature has been mostly focused on the lodgin industry. Deng, Veiga and Wiper (2019) employed a cross-sectional SFA to explain the efficiency of Spanish and Portuguese hotel chains in 2014, concluding that it is more efficient to invest in a few big hotels rather than several small ones. Some authors have focused on smaller regions instead of the whole country. For instance, Martínez-Roget and

Rodríguez-González (2006) employed SFA to analyze efficiency and total factor productivity (TFP) in the Galician lodging sector. The accommodation sector in the Canary Islands has also been a focus of study: for example Pérez-Rodríguez and Acosta-González (2007) analyzed the evolution of cost efficiency for Gran Canaria's hotels and tourist apartments during the period 1991 and 2002 with an SFA. The results showed that the sector grew during this period by making cost reductions and productivity gains. However, the authors found decreasing returns to scale for most units in the sample. Arbelo-Pérez, Pérez-Gómez and Arbelo (2019) also studied the case of the Canary Islands, but focused on the impact of the tourism packages by using a panel-data SFA; and concluded that hotels with tourist packages are less efficient. Further, these authors argue that even when the hotels are cost-efficient, efficiency in terms of profits is lower. Efficiency analysis can be used to assess labour productivity. An example of this can be found in Cordero and Tzeremes (2018) and Tzeremes (2020) who estimated labour productivity for Spanish hotels in the two Spanish archipelagos using DEA approaches. Additionally, Inchausti-Sintes, Pérez-Granja and Morales-Mohamed (2020) used an SFA approach to estimate labour productivity and introduced it into a CGE model to calculate its economic impact. Efficiency analysis has also been applied to tourism fairs, showing significant differences among the exhibitors depending on the sector and type (Alberca-Oliver, Rodríguez-Oromendia & Parte-Esteban, 2015).

At macro level, the analysis of Spain has concentrated on Autonomous regions (NUTS II), where determinants of efficiency have been the main question for these studies. For instance, Benito, Solana and López (2014) estimate the effect on the efficiency of nine tourism attractions and used the efficiency scores as a proxy of tourism competitiveness. This paper concluded that Catalonia, the Canary Islands, Andalucia, the Balearic Islands and the Basque Country are the most efficient NUTS II regions. Similarly, Solana Ibañez, Caravaca-Garratón and Para-González (2016a) also studied the evolution of Spanish tourism productivity. The results showed that the Balearic Islands, Canary Islands and Murcia are on the frontier, while productivity remained almost constant during the period of analysis. The role of environmental variables has also been analyzed by Sellers-Rubio and Casado-Díaz, with the authors finding a high level of inefficiency for the Spanish Autonomous Regions.

Stochastic Frontiers in Tourism

Compared with DEA, SFA has been less used in the tourism literature. Since Barros (2004), who focused on the efficiency of publicly-owned hotels in Portugal, few studies have applied the methodology (see for example: Barros, 2006; Chen, 2007; Cracolici, Nijkamp & Rietveld, 2008; Pavlyuk, 2011; Assaf, Oh & Tsionas, 2017; or Wu, Cheng & Liao, 2019). Most of these studies have focused on the models proposed by Battese and Coelli (1992, 1995) (see for example: Barros, 2004 and 2006; Roget & Rodríguez-González, 2006; Pérez-Rodríguez & Acosta-González, 2007; Chen, 2007; Zhou, Xu & Lee, 2019; Ripoll-Zarraga & Raya, 2020). The former allows for a constant growth rate in efficiency while the latter allows for the inclusion of determinants in inefficiency. Other authors have opted for different approaches, such as Fernández, Coto-Millán and Díaz Medina (2018), who used the True Random Effects model (Greene, 2005). This approach extends the classic random effect model by disentangling panel heterogeneity from inefficiency by assuming that inefficiency is only time-varying. The Bayesian models represent a recent alternative approach in SFA (Assaf, Oh & Tsionas, 2017; or Deng, Veiga & Wiper, 2019). Briefly, this modelization is based on the Bayes rule to simultaneously tackle data (likelihood) and previous knowledge about certain parameters (prior distribution) to obtain the estimates of the parameters of the inefficiency term (posterior distribution). Finally, Lien, Kumbhakar and Alem (2018) provide one of the most recent developments in SFA. This model deals with some of the caveats identified by Assaf and Tsionas (2019), such as endogeneity and heterogeneity issues. The model also disentangles inefficiency in two components: time variant inefficiency and time invariant inefficiency (see section 4 for more details).

[Table 1 about here]

3. Case-Study

Spain has 50 provinces, with many characterized by sharp differences in their productive-mix. On the one hand, the country has historical northern industrial-led provinces such as Álava, Vizcaya and Guipúzcoa in the Basque Country and Navarre. According to the Spanish Statistical Institute (INE), in 2018, the industrial and service sectors in both regions averaged around 20% and 46%, and 24% and 40% of total GDP, respectively. On the other hand, tourism-led economies include the two Canarian provinces; the Balearic province (formally known as the Canary and the Balearic Islands, respectively); and some Mediterranean provinces. For instance, in the case of the Canaries and the Balearic Islands, the industrial and services sectors averaged 2.8% and 60%, and 2.4% and 64.2%, respectively.

The industrial and service sector provide a useful but limited approach to measure the impact of tourism specialization and potential spillover effect from adjacent and more productive sectors. In order to narrow the approach to tourism specialization, two different variables were taken into consideration. On the one hand, the number of tourist attractions (which includes natural parks, monuments, museums, amusement parks, zoos and aquatic parks) was included as a proxy of tourism specialization. On the other, the ratio of 4 and 5 star hotels over total accommodation was used as a proxy of the quality of the accommodation supply. Both variables are fixed at province level due to the fact that they were obtained from Tripadvisor in 2020. Additionally, in order to account not only for the industrial weight, but also for its level of competitiveness the ratio of exports of goods over the total gross value added (GVA) was used as proxy of spillover effects from more productive activities. This variable measures the level of international competitiveness of the tradable goods produced in each province.

Figure 1 is consistent with the aforementioned description by showing the geographical pattern enhanced by specialization in services. Clearly, except for Madrid, all service-based provinces are located in the Mediterranean and both archipelagos which, at the same time, have the main ‘sun and beach’ destinations in the country. Consequently, these provinces also concentrated most of the tourism hotel supply.

[Figure 1 about here]

On the other hand, tourism seasonality is also present in all provinces, with the clear exception of the Canarian archipelago and, to a lesser extent, the cities of Madrid and Barcelona (Spain’s biggest cities) (Figure 2). The Balearic archipelago has the greatest seasonality.

[Figure 2 about here]

Table 2 summarizes the main statistics of the variables used in the model. The table shows that there is significant variance in all the reported variables. These differences by provinces can be more easily identified by analyzing the minimum and maximum values of the variables. For instance, on the one hand, provinces such as Palencia, Soria, Zamora or Guadalajara are all located in what is known as “hollow Spain”, which show overnight stays close to the minimum. On the other hand, tourism-led provinces such as the two archipelagos, the mainland Mediterranean destinations and Barcelona and Madrid, concentrate the majority of tourism overnight stays. Likewise, as might be expected, these latter provinces also report the highest values in terms of “average annual hotel beds” and “average annual labour at hotels”.

[Table 2 about here]

Again, there are significant differences among the Spanish provinces when analyzing seasonal concentration (measured with a seasonal Gini index by provinces (see Fernández-Moralez et al. 2016 for further details). In this sense, the Canary Islands, Madrid and Barcelona show the lowest seasonality. In the case of the Canaries seasonality is technically inexistent, while the Balearic Islands, Girona and Tarragona are the most seasonal. Finally, the share of industry and services also show significant differences, as already noted at the beginning of this section.

4. Methodology

Stochastic Frontier Analysis

Stochastic Frontier Analysis (SFA) is a parametric approach which estimates efficiency by computing a stochastic production function given the inputs, or a stochastic frontier cost function given the outputs. The idea behind this kind of model is that the error term can be disentangled into two components: $\epsilon_{it} = v_{it} + u_{it}$ where ϵ_{it} is the classic error term, v_{it} is the unobserved random component and u_{it} denotes the inefficiency component. As a result, most of the contributions in this field focus on providing an alternative way of modeling the inefficiency component.

The most basic SFA model in panel data takes the form of a simple fixed or random effect where the panel specific variable is related to the inefficiency term. In these models, there is no assumption about the behaviour of the inefficiency. This approach was introduced by Schmidt and Sickles (1984). However, these models assumed that the inefficiency was constant over time, which can be a restrictive assumption, especially if the time horizon is big enough. For these reasons Cornwell, Schmidt and Sickles (1990) propose a model where the inefficiency was allowed to change over time. Since then, different ways of dealing with the evolution of inefficiency have been proposed by Kumbhakar (1990), Battese and Coelli (1992) and Kumbhakar and Wang (2004). Battese and Coelli (1995) propose a model were the inefficiency term can be modelled with explanatory variables which are allowed to be either time-variant or time-invariant.

Nonetheless, none of these models can separate panel heterogeneity from the inefficiency. For this reason, Greene (2005) proposed both the True Fixed Effects (TFE) and the True Random-effects models (TRE), which allowed researchers to disentangle panel heterogeneity from the inefficiency term. On the other hand, Greene (2005) assumed that the persistent parameter is part of panel heterogeneity and therefore, the inefficiency can only be time-variant. In sum, the aforementioned models cannot fulfill the three conditions at the same time (i.e. they cannot distinguish between permanent and transient inefficiency, and they separate firm/individual heterogeneity from the permanent inefficiency). In this context, Kumbhakar, Lien and Hardaker (2014) extended the TRE model into a four components model known as the General True Random

Effects model (GTRE) (Tsionas and Kumbhakar, 2014). The latter allows the measurement of these three components by generating a three step procedure. Colombi, Kumbhakar, Martini and Vittadini (2014) proves that this model can also be estimated in a single step by using the Maximum Likelihood function. However, this approach is hard to apply in practice (Filippini and Greene, 2016). Finally, Lien, Kumbhakar and Alem (2018) enrich the GTRE model by allowing the inclusion of determinants into the inefficiency terms. Moreover, this approach deals with endogeneity, which has been an omitted issue in most SFA analysis.

The General True Random Effects models with determinants of persistent inefficiency.

Lien, Kumbhakar and Alem (2018) introduce and apply the model to Norwegian crop-producing farms. One of the particularities of this approach is dealing with endogeneity. In tourism, decisions about the level of inputs (labour and available beds) are not independent of expected demand in each period, which can be thought of in terms of peak and off-peak seasons. This implies that inputs and outputs are economically endogenous. Econometrically, this generates problems by breaking the assumption of independence between explanatory variables, with inefficiency and the error term leading to biased estimates. To deal with endogeneity the model assumes that producers are maximizing the return to outlay (RO), which is defined as total revenue divided by total cost. Under this assumption the production function is homogeneous of degree one. This homogeneity allows the model to be rewritten in relative terms of one factor; and this reformulation allows for the inputs to be uncorrelated with the error and inefficiency terms (those interested in the mathematical proof are referred to Lien, Kumbhakar and Alem, 2018).

The panel data efficiency model can be written as follows:

$$ln\tilde{y}_{it} = \beta_0 + \sum_{j=2} \beta_j ln\tilde{x}_{it} + \beta_t t + \beta_{t2} t^2 + \mu_i - \eta_i(z_i) + v_{it} - u_{it} \quad (1)$$

Were $\tilde{y}_{it} = \frac{y_{it}}{x_{1it}}$, $\tilde{x}_{it} = \frac{x_{jit}}{x_{1it}}$ is a vector of inputs, t is a time trend, which also appears squared to capture the non-lineal trend. The variable μ_i represents panel heterogeneity, v_{it} is the error term, $\eta_i(z_i)$ is a non-negative value representing persistent (long run) inefficiency and depends on a vector z_i of determinants. Lastly, u_{it} is a non-negative value denoting the time-varying (short run) inefficiency.

The panel-effect μ_i is assumed to be random and i.i.d N($(0, \sigma_\mu^2)$. The random shocks v_{it} are also assumed i.i.d. $N(0, \sigma_v^2)$. $E\left(\eta_i(z_i)\right) = g(z_i) \geq 0$ and $E(u_{it}) \geq 0$ depending on the assumptions on the distribution of u_{it}. Equation (1) can be rewritten as:

$$ln\tilde{y}_{it} = [\beta_0 - g(z_i) - E(u_{it})] + \beta' ln\tilde{x}_{it} + [\mu_i - (\eta_i(z_i) - g(z_i))] + [v_{it} - (u_{it} - E(u_{it}))]$$

$$ln\tilde{y}_{it} = h(z_i) + \beta' ln\tilde{x}_{it} + a_i(z_i) + \varepsilon_{it} \quad (2)$$

Where $h(z_i) = \beta_0 - g(z_i) - E(u_{it})$, $a_i(z_i) = \mu_i - (\eta_i(z_i) - g(z_i))$ and $\varepsilon_{it} = v_{it} - (u_{it} - E(u_{it}))$. It can be seen that, $E(a_i(z_i))$ and $E(\varepsilon_{it})$ are equal to 0. Thus the model described in (2) is the partially linear version of the random effects panel data model (see Robinson, 1988); so it cannot be estimated by a linear approach due to consistency problems with the parameters. Robinson (1988) suggests estimating the model by taking the conditional expectation in (2), respecting z_i.

$$E\left(ln\tilde{y}_{it} \middle| z_i\right) = E(h(z_i) + \beta' ln\tilde{x}_{it} + a_i(z_i) + \varepsilon_{it})|z_i$$

$E\left(ln\tilde{y}_{it}\middle|z_i\right) = E(h(z_i)|z_i + \beta' E(ln\tilde{x}_{it}|z_i) + E(a_i(z_i)|z_i) + E(\varepsilon_{it}|z_i)$

Knowing that $E(a_i(z_i)|z_i)$ and $E(\varepsilon_{it}|z_i)$ are equal to 0 then:

$$E\left(ln\tilde{y}_{it}\middle|z_i\right) = h(z_i) + \beta' E(ln\tilde{x}_{it}|z_i) \qquad (3)$$

We can now subtract (3) to (2) to obtain:

$$ln\tilde{y}_{it} - E(ln\tilde{y}_{it}|z_i) = \beta'[ln\tilde{x}_{it} - E(ln\tilde{x}_{it}|z_i)] + a_i(z_i) + \varepsilon_{it} \qquad (4)$$

The conditional means $E\left(ln\tilde{y}_{it}\middle|z_i\right)$ and $E(ln\tilde{x}_{it}|z_i)$ are estimated using a non-parametric regression. After estimating the conditional expectations, we can rewrite (4) as:

$$y^*_{it} = \beta' x^*_{it} + a_i(z_i) + \varepsilon_{it} \qquad (5)$$

Where $y^*_{it} = ln\tilde{y}_{it} - E\left(ln\tilde{y}_{it}\middle|z_i\right)$ and $x^*_{it} = ln\tilde{x}_{it} - E(ln\tilde{x}_{it}|z_i)$. The model in (5) is a linear random effects panel data model. This model can be estimated by linear regression obtaining consistent estimates of the parameters. It also gives predicted values of $a_i(z_i)$ and ε_{it}, which can be use in a second and third step. This is precisely the model applied in our case study. More specifically, the model can be summarized as follows:

$$ln\tilde{y}_{it} = \beta_0 + \beta_1 ln\tilde{k}_{it} + \beta_2 ln\tilde{k}^2_{it} + \beta_3 t + \beta_4 t^2 + \beta_5 ln\tilde{k}_{it} t + \beta_6 financial\ crisis +$$
$$\beta_6 sovereign\ debt\ crisis + \mu_i + v_{it} + \eta_i(z_i) + u_{it} \qquad (6)$$

Where:

v_{it} is i.i.d. $N(0, \sigma^2_v)$

u_{it} is i.i.d $N^+(0, \sigma^2_u)$

μ_i is i.i.d. N($\left(0, \sigma^2_\mu\right)$

η_i is i.i.d $N^+\left(0, \sigma^2_\eta(z_i)\right) = N^+(0, exp(\delta_{i0} + \delta'_{i1} z_i))$.

As shown in equation (1), the model adopts a translog production function. The endogenous variable ($\tilde{y}_{it}$) represents the ratio of overnights over labour in the hotel industry. $\tilde{k}_{it}$ is the ratio of the hotel beds (a proxy of capital in the hotel sector) over labour in the hotel industry. Variable t is a time trend. The crisis effect has been split into $financial\ crisis$ and $sovereign\ debt\ crisis$ as suggested in Tzeremes (2019) and is represented by dummy variables accounting for each economic crisis. The variable μ_i represents panel heterogeneity, v_{it} is the error term and $\eta_i(z_i)$ is a non-negative value representing the persistent inefficiency and depends on a vector z_i of determinants. In this study, the determinants included consist in the average share of industry, the average share of services and the average seasonal concentration index. Finally, u_{it} is a non-negative value representing the time-varying inefficiency. Likewise, equation (6) can be transformed using the aforementioned method resulting in:

$$ln\tilde{y}_{it} = \beta_1 ln\tilde{k}_{it} + \beta_2 ln\tilde{k}_{it}^2 + \beta_3 t + \beta_4 t^2 + \beta_5 ln\tilde{k}_{it} t + \beta_6 financial\ crisis + \beta_6 sovereign\ debt\ crisis + a_i^*(z_i) + \varepsilon_{it}^* \quad (7)$$

At the first stage, as suggested by Lien, Kumbhakar and Alem (2018), equation (7) is estimated by random effects, obtaining the β coefficients and the predicted values of α_i and ε_{it}. On the second and third steps, a pooled Stochastic Frontier is conducted over the predicted values of ε_{it}^* and $a_i^*(z_i)$ respectively. The transient and persistent inefficiency (and efficiency) are estimated using the Jondrow, Lovell, Materov and Schmidt (1982) and Battese and Coelli (1988) procedures, using the output of the second and third step respectively. These procedures are used in order to obtain the technical inefficiency (Jondrow et al.,1982) and the technical efficiency (Battese and Coelli, 1988).

Results

Two models were estimated in STATA 14. Table 3 shows the main results of all stages. On the one hand, model 1 refers to the model where the share of services and industry were introduced as determinants of the persistent inefficiency. On the other, model 2 reports the results when using more accurate proxy variables of tourism specialization, and adjacent sectoral competitiveness. More precisely, this model uses the share of exports of goods over the gross value added of each province as a proxy of the latter. Moreover, it accounts for tourism specialization by including the number of tourism attractions of each province; and the share of 4 and 5 star hotels over the total accommodation to account for the quality of the supply. It can be seen that both models provide similar results in terms of coefficients and efficiency scores. Additionally, the economic crisis had a long-lasting effect on the Spanish economy. In this sense, both variables are significant. However, according to the coefficients, the sovereign crisis is confirmed as having a greater impact than the financial crisis.

The analysis of the determinants of the inefficiency term is particularly interesting, as it is modeled as part of the persistent inefficiency. The results show that all the determinants prove to be significant and with a negative sign in both models. This means that an increase in these variables will reduce inefficiency. However, it is interesting to analyze each of them independently. Firstly, we have seasonal concentration. This variable captures how tourist overnight stays are concentrated during the year. Intuitively, as suggested in the introduction, one might presume that provinces with a more homogeneously distributed flow of tourists should be more efficient because they do not have marked peak and off-peak seasons. However, at first glance, our findings show the opposite result (i.e. an increase in seasonal concentration aids efficiency). This might be explained by the fact that, in highly seasonal markets, such as the Balearic

Islands, most hoteliers close their establishments during off-peak seasons (losing fixed costs). On the other hand, in places where tourist flows are more stable, the majority of supply is kept open, so there are some losses in efficiency. For instance, in 2016, the bed supply in the Balearic Islands varies between 18,732 and 347,795 in the off-peak and peak month, respectively. The off-peak month represents 5% of the peak month. Conversely, the bed supply in the Canary Islands is 141,381 in the off-peak month and 149,090 in the peak month.

[Table 3 about here]

Analyzing Model 1, it can be seen that, firstly, the share that the industry has also contributes to reducing the inefficiency (-7.742). This means that an increase in the share of the industry in the economy reduces the inefficiency in these provinces. Secondly, the share of services has a similar interpretation to the previous one, but with the latter showing a more intense effect than the former (-50.456) (i.e. service-based economies, such as those that are tourism-led, reduce inefficiencies in the hotel sector more effectively than those that are industry-led).

Analyzing Model 2, the share of exports of goods over the GVA shows a negative and significant sign (-0.618), meaning that provinces producing more competitive goods are more efficient, while enhancing a more efficient hotel industry. Focusing on tourism determinants, it can be seen that both the number of tourism attractions per km^2 (-72.213) and the ratio of 4 and 5 star hotels (-11.175) help reduce the inefficiency. Nevertheless, the difference in magnitude of both variables is significant. In this sense, an increase in the quality of the tourism supply reduces inefficiency. However, 'tourism attractiveness' seems to have greater impact. Finally, none of the models could capture any significant transient inefficiency, while the persistent efficiency averages 0.862 in Model 1 and 0.849 in Model 2. This means that the efficiency of the Spanish provinces is high, however, there is 'permanent inefficiency' in the hotel sector of the Spanish provinces, which is intrinsic to its productive-mix.

[Table 4 about here]

As shown above, the GTRE model disentangles the inefficiency into transient and persistent inefficiency. This allows us to measure which part of the inefficiency is intrinsic to the provinces and which part varies through time. In order to avoid presenting over-complex tables and graphics, Table 4 shows the efficiency scores, but aggregated by clusters. These clusters were created by 'K-means', taking into account the following variables: the number of attractions (natural parks, monuments, museums, amusement parks, zoos and aquatic parks); the ratio of 4 and 5 star hotels over the total; the average share of the industrial sector; the average share of the service sector; the average ratio of foreign tourists over the total; the average number of overnights; the average number of hotel beds; the average number of workers in hotels; and the average length of stay. Additionally, from now on, as both models reported similar results, only results from Model 2 are shown.

Table 4 shows the average and the standard deviation of each efficiency score in each cluster. As can be appreciated when analyzing this table and figure, the differences in efficiency are minimal among tourism-led economies (the Balearic Islands, the Canary Islands and the Mediterranean destinations) and the Big Cities, while they differ among themselves and the other provinces. The comparison between the archipelagos is especially relevant taking into account that, as shown in Figure 2, both the Balearic and the Canary Islands show the highest and lowest seasonality, respectively. On the other hand, the Canary Islands show a relatively smaller efficiency score than the Balearic Islands; providing evidence that all-year-round tourism does not necessarily mean it is the most efficient.

Analyzing the persistent efficiency, a clear gap can be seen between tourism-led regions and the other regions, with the exception of the “big cities” (the provinces of Madrid and Barcelona). These results are in line with tourism specialization being the key driver of efficiency in the hotel sector. The score of the Balearic Islands is especially significant as it is, according to our findings, the most efficient region in Spain; while, at the same time, it is the most seasonal. This result is aligned with the aforementioned role of seasonal concentration as a determinant of efficiency. Nevertheless, the level of persistent efficiency in Spain’s big cities reaches a level similar to that obtained in the Balearic archipelago. This can be explained by the size of these economies and their more diversified productive-mix, which is significantly higher than any other province in the country. Meanwhile, at the same time, they attract significant flows of tourism throughout the year. The cluster composed of mainland Mediterranean sun and beach destinations and the Canary Islands enjoys very similar efficiency scores. Finally, the provinces located in the centre and northern part of the mainland show the worst efficiency scores. These kinds of provinces do not show a clear pattern in terms of economic specialization (services or industry).

[Figure 3 about here]

Figure 3 facilitates the visualization of the efficiency by provinces according to their score; where an intense red means that the province is more efficient. As shown, the Mediterranean coastal provinces and both archipelagos dominate the efficiency score. Nevertheless, it is notable that the efficiency steadily reduces on the Mediterranean coast as soon as we move from north to south. The map helps to highlight the suitability of adopting a NUTS III approach. For instance, the autonomous community of Andalusia (NUTS II) is formed by a varying group of provinces such as Cadiz, Córdoba, Jaen, Huelva, Almeria, Malaga, Granada and Seville (south of Spain). Adopting a NUTS III approach allows us to distinguish tourism-oriented provinces such as Malaga, from other non-tourism-oriented ones such as Jaen or Córdoba, enriching the analysis. The provinces with the lighter colour are mainly concentrated in the centre of the country, surrounding the province of Madrid. As previously noted, on average none of them receive significant tourism flows or show a highly developed industrial or services sector.

Lastly, given the lack of transient inefficiency, different SFA models were obtained to compare with the approach used in this paper. All the SFA obtained used the variables of model 2. Three different panel data SFA models that allow for using determinants of inefficiency are shown in Table 5. The first model refers to the GTRE model used before. The second is the True Random Effects model (TRE) (Greene, 2005), and the third one corresponds to the Battese and Coelli (1995) (BC95) model. These two additional models are widely used in the panel data literature (Kumbhakar et al., 2014). The main methodological differences among the models are addressed when analyzing the results below.

[Table 5 about here]

It can be seen that the coefficients of the stochastic frontier are relatively similar between the GTRE and the TRE, given that the former is an extension of the latter; while the parameters differ significantly with the BC95 model regarding to the capital/labour ratio factor. However, it remains similar for the other coefficients, but with a higher standard error.

Analyzing the determinants of efficiency variance, it can be seen that the main source of divergence is the ratio of 4 and 5 stars hotel that is not significant neither for the TRE nor for the BC95. It should be noted that the TRE cannot capture the inefficiency given the fact that, as showed in the LKA model, this inefficiency is persistent. The TRE model can identify the transient inefficiency, but cannot differentiate the persistent one from the firm heterogeneity resulting in an infraestimation of the inefficiency. The BC95 model shows a similar level of inefficiency to the LKA model. However, it should be noted that this model is generated in a two-step approach where the inefficiency is regressed on the determinants. Thus, this model does not allow us to differenciate between transient (short run effects) and persistent (long run effects) efficiency, nor panel heterogeneity. Thus, it is more difficult to properly differentiate which determinants of inefficiency are structural and which are manageable in the short run.

6. Conclusions

This paper analyzes the extent to which the goodness of economic specialization detected in other economies and sectors is also confirmed in the hotel sector, in tourism service-based economies. As described in the introduction, the particularities of these kinds of economies are: seasonality; lack of economic diversification; and limitations in the outsourcing of services; which draws into question this result in these kinds of economies. Likewise, industrial-led economies may also achieve higher spillover effects by transferring "know-how", knowledge or technology from industrial or adjacent more productive activities to services fostering efficiency. Finally, the analysis is based on applying the recent model developed by Lien, Kumbhakar and Alem (2018), which allows determinants in the inefficiency term and deals with endogeneity; an omitted issue in most SFA analysis. In this sense, the efficiency is modelled by controlling for both service and industrial specialization, and seasonality. Finally, the model also distinguishes three components: heterogeneity, and permanent and transient inefficiency components.

The results confirm that specialization improves efficiency in the hotel sector. More precisely, it was approached in different ways. On the one hand, the number of tourist attractions per km^2 and the ratio of 4 and 5 star hotels over total accommodation was included as a proxy of tourism specialization. On the other, we took into account potential spillover effects from other more productive sectors, by capturing the level of competitiveness of each province economy using the ratio of exports of goods over the total gross value added (GVA). This variable measures the level of international competitiveness of the tradable goods produced in each province.

The study has also identified another novel result by showing that lower seasonality does not necessarily mean greater efficiency. In other words, seasonality is not detrimental to efficiency in the hotel sector. In this sense, the Balearic Islands, which show both the highest seasonality and the highest efficiency score. On the other hand, the Canaries show the second-highest score, which is very similar to the former, but has the lowest seasonality of all Spanish provinces. However, taking into account the three determinants of the inefficiency term: seasonality, and industrial and service specialization,he latter has the strongest impact of the three in terms of reducing inefficiency; confirming the goodness of tourism specialization by enhancing efficiency in the hotel sector. As a result, tourism-led provinces were the most efficient Spanish economies in this sector during the period 2001-2016.

Finally, the analysis also allows us to emphasize specific results from an economic policy perspective. Firstly, the greater efficiency in the hotel sector in tourism-led provinces implies, together with the goodness of its specialization, an important source for productivity improvement in a sector with serious difficulties in making progress in this aspect (the industry, the main source of productivity gains in advanced economies, represents a marginal share in most tourism-led economies), allowing for stronger competitiveness and an increase in salaries in the long term. Secondly, seasonality has always been a political matter of concern in tourism-led economies because of the high temporality of employment. However, taking into account the positive effect of the former in efficiency, future employment policies should be redefined to cope with specific labour training in off-peak seasons to keep maintaining and improving the efficiency in the sector, that, similar to the previous point, may also enhance productivity gains and higher salaries in the long term. Finally, these effects are especially relevant for the Spanish developed tourism destinations in order to compete with cheaper emerging alternatives.

Likewise, the importance of persistent inefficiency reveals the structural nature of the inefficiency of the hotel accommodation sector that, together with the importance of tourism attractiveness and quality supply, is highlighting the need to implement long-term policies aimed at the maintenance of the former and the improvement of the latter to ensure the competitiveness and revitalization of mature tourism destinations in the country.

References

Alberca-Oliver, P., Rodríguez-Oromendía, A., & Parte-Esteban, L. (2015). Measuring the efficiency of trade shows: A Spanish case study. *Tourism Management*, *47*, 127–137. https://doi.org/10.1016/j.tourman.2014.09.007

Alemayehu, F. K., & Tveteraas, S. L. (2019). Long-run labour flexibility in hospitality: A dynamic common correlated effects approach. *Tourism Economics*, 1354816619864802. https://doi.org/10.1177/1354816619864802

Arbelo-Pérez, M., Pérez-Gómez, P., & Arbelo, A. (2019). Impact of all-inclusive packages on hotel efficiency. *Current Issues in Tourism*, *22*(8), 905–920. https://doi.org/10.1080/13683500.2017.1323850

Assaf, A. G., & Josiassen, A. (2016). Frontier Analysis: A State-of-the-Art Review and Meta-Analysis. *Journal of Travel Research*, *55*(5), 612–627. https://doi.org/10.1177/0047287515569776

Assaf, A. G., & Tsionas, M. G. (2019). A review of research into performance modeling in tourism research - Launching the Annals of Tourism Research curated collection on performance modeling in tourism research. In *Annals of Tourism Research* (Vol. 76, pp. 266–277). https://doi.org/10.1016/j.annals.2019.04.010

Assaf, A. G., Oh, H., & Tsionas, M. (2017). Bayesian Approach for the Measurement of Tourism Performance: A Case of Stochastic Frontier Models. *Journal of Travel Research*, *56*(2), 172–186. https://doi.org/10.1177/0047287516633528

Barros, C. P. (2004). A stochastic cost frontier in the Portuguese hotel industry. *Tourism Economics*, *10*(2), 177–192. https://doi.org/10.5367/000000004323142416

Barros, C. P. (2006). Analysing the rate of technical change in the Portuguese hotel industry. *Tourism Economics*, *12*(3), 325–346. https://doi.org/10.5367/000000006778493673

Battese, G. E., & Coelli, T. J. (1988). Prediction of firm level technical inefficiencies with a generalized frontier production function. *Journal of econometrics*, *38*, 387-399.

Battese, G. E., & Coelli, T. J. (1992). Frontier production functions, technical efficiency and panel data: with application to paddy farmers in India. *Journal of productivity analysis*, *3*(1-2), 153-169. https://doi.org/10.1007/BF00158774

Battese, G. E., & Coelli, T. J. (1995). A model for technical inefficiency effects in a stochastic frontier production function for panel data. *Empirical economics*, *20*(2), 325-332. https://doi.org/10.1007/BF01205442

Ben Aissa, S., & Goaied, M. (2016). Determinants of Tunisian hotel profitability: The role of managerial efficiency. *Tourism Management*, *52*, 478–487. https://doi.org/10.1016/j.tourman.2015.07.015

Benito, B., Solana, J., & López, P. (2014). Determinants of Spanish regions' tourism performance: A two-stage, double-bootstrap data envelopment analysis. *Tourism Economics*, *20*(5), 987–1012. https://doi.org/10.5367/te.2013.0327

Bougheas, S., Demetriades, P.O. & Manuneas, T.P. (2000) Infrastructure, specialization, and economic growth, *Canadian Journal of Economics, 33*(2), 506- 622. https://doi.org/10.1111/0008-4085.00026

Brau, R., Lanza, A., & Pigliaru, F. (2007). How fast are small tourism countries growing? Evidence from the data for 1980–2003. *Tourism Economics*, *13*(4), 603-613. https://doi.org/10.5367/000000007782696104

Capo, J., Font, A. R., & Nadal, J. R. (2007). Dutch disease in tourism economies: Evidence from the Balearics and the Canary Islands. *Journal of sustainable Tourism*, *15*(6), 615-627. https://doi.org/10.2167/jost698.0

Chang, Y. T., Lee, S., & Park, H. (Kevin). (2017). Efficiency analysis of major cruise lines. *Tourism Management*, *58*, 78–88. https://doi.org/10.1016/j.tourman.2016.10.012

Chao, C. C., Hazari, B. R., Laffargue, J. P., Sgro, P. M., & Eden, S. H. (2006). Tourism, Dutch disease and welfare in an open dynamic economy. *The Japanese Economic Review*, *57*(4), 501-515.

Chen, C. F. (2007). Applying the stochastic frontier approach to measure hotel managerial efficiency in Taiwan. *Tourism Management*, *28*(3), 696–702. https://doi.org/10.1016/j.tourman.2006.04.023

Colombi, R., Kumbhakar, S. C., Martini, G., & Vittadini, G. (2014). Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. *Journal of Productivity Analysis*, *42*(2), 123-136. https://doi.org/10.1007/s11123-014-0386-y

Cordero, J. M., & Tzeremes, N. G. (2018). Financial crisis and hotels' labour productivity growth: evidence from Spanish islands. *Applied Economics Letters*, 25(19), 1376-1382. https://doi.org/10.1080/13504851.2017.1420884

Cornwell, C., Schmidt, P., & Sickles, R. (1990). Production frontiers with cross-sectional and time-series variation in efficiency levels. *Journal of Econometrics*, 46 (1-2), 185-200.

Cracolici, M. F., Nijkamp, P., & Rietveld, P. (2008). Assessment of tourism competitiveness by analysing destination efficiency. *Tourism economics*, *14*(2), 325-342. https://doi.org/10.5367%2F000000008784460427

Cuccia, T., Guccio, C., & Rizzo, I. (2017). UNESCO sites and performance trend of Italian regional tourism destinations: A two-stage DEA window analysis with spatial interaction. *Tourism Economics*, *23*(2), 316–342. https://doi.org/10.1177/1354816616656266

De Vita, G., & Kyaw, K. S. (2017). Tourism specialization, absorptive capacity, and economic growth. *Journal of Travel Research*, *56*(4), 423-435. https://doi.org/10.1177/0047287516650042

Deng, Q., Gu, X., Law, R., & Lian, Z. (2020). A comparative study for determinants of gaming performance in Macao and Las Vegas. *Tourism Management*, *77*, 103964. https://doi.org/10.1016/j.tourman.2019.103964

Deng, Y., Veiga, H., & Wiper, M. P. (2019). Efficiency evaluation of hotel chains: a Spanish case study. *SERIEs*, *10*(2), 115–139. https://doi.org/10.1007/s13209-019-0188-6

Eugenio-Martin, J. L., Cazorla-Artiles, J. M., & González-Martel, C. (2019). On the determinants of Airbnb location and its spatial distribution. *Tourism Economics*, *25*(8), 1224-1244. https://doi.org/10.1177/1354816618825415

Feenstra, R. C., & Taylor, A. M. (2011). *International trade*. New York: Worth Publishers.

Fernández, X. L., Coto-Millán, P., & Díaz-Medina, B. (2018). The impact of tourism on airport efficiency: The Spanish case. *Utilities Policy*, *55*, 52-58. https://doi.org/10.1016/j.jup.2018.09.002

Fernández-Morales, A., Cisneros-Martínez, J. D., & McCabe, S. (2016). Seasonal concentration of tourism demand: Decomposition analysis and marketing implications. *Tourism Management*, 56, 172-190. https://doi.org/10.1016/j.tourman.2016.04.004

Filippini, M., & Greene, W. (2016). Persistent and transient productive inefficiency: a maximum simulated likelihood approach. *Journal of Productivity Analysis*, *45*(2), 187-196. https://doi.org/10.1007/s11123-015-0446-y

Fixler, D. J., & Siegel, D. (1999). Outsourcing and productivity growth in services. *Structural change and economic dynamics*, *10*(2), 177-194. https://doi.org/10.1016/S0954-349X(98)00048-4

Greene, W. (2005). Fixed and random effects in stochastic frontier models. *Journal of productivity analysis*, *23*(1), 7-32.https://doi.org/10.1007/s11123-004-8545-1

Huang, C. wei. (2018). Assessing the performance of tourism supply chains by using the hybrid network data envelopment analysis model. *Tourism Management*, *65*, 303–316. https://doi.org/10.1016/j.tourman.2017.10.013

Inchausti-Sintes, F. (2019). A tourism growth model. *Tourism Economics*, 1354816619840096. https://doi.org/10.1177/1354816619840096

Inchausti-Sintes, F., Pérez-Granja, U., & Morales-Mohamed, J. J. (2020). Analysing labour productivity and its economic consequences in the two Spanish tourist archipelagos. Tourism Economics (in press). https://doi.org/10.1177%2F1354816620917865

Jondrow, J., Lovell, C. K., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of econometrics*, *19*(2-3), 233-238.

Karakitsiou, A., Kourgiantakis, M., Mavrommati, A., & Migdalas, A. (2018). Regional efficiency evaluation by input-oriented data envelopment analysis of hotel and restaurant sector. *Operational Research*, 1–18. https://doi.org/10.1007/s12351-018-0406-1

Krugman, P., & Venables, A. J. (1996). Integration, specialization, and adjustment. *European economic review*, *40*(3-5), 959-967. https://doi.org/10.1016/0014-2921(95)00104-2

Kularatne, T., Wilson, C., Månsson, J., Hoang, V., & Lee, B. (2019). Do environmentally sustainable practices make hotels more efficient? A study of major hotels in Sri Lanka. *Tourism Management*, *71*, 213–225. https://doi.org/10.1016/j.tourman.2018.09.009

Kumbhakar, S. C. (1990). Production frontiers, panel data, and time-varying technical inefficiency. *Journal of econometrics*, *46*(1-2), 201-211. https://doi.org/10.1016/0304-4076(90)90055-X

Kumbhakar, S. C., & Wang, H. J. (2004). Estimation of growth convergence using a stochastic production frontier approach. Economics Letters, 88(3), 300-305. https://doi.org/10.1016/j.econlet.2005.01.023

Kumbhakar, S. C., Lien, G., & Hardaker, J. B. (2014). Technical efficiency in competing panel data models: A study of Norwegian grain farming. *Journal of Productivity Analysis*, *41*(2), 321–337. https://doi.org/10.1007/s11123-012-0303-1

Lee, J. (2011). Export specialization and economic growth around the world. *Economic systems*, *35*(1), 45-63. https://doi.org/10.1016/j.ecosys.2010.11.002

Lien, G., Kumbhakar, S. C., & Alem, H. (2018). Endogeneity, heterogeneity, and determinants of inefficiency in Norwegian crop-producing farms. *International Journal of Production Economics*, *201*, 53–61. https://doi.org/10.1016/j.ijpe.2018.04.023

Nowak, J. J., & Sahli, M. (2007). Coastal tourism and 'Dutch disease'in a small island economy. *Tourism Economics*, *13*(1), 49-65. https://doi.org/10.5367/000000007779784452

Parrilla, J. C., Font, A. R., & Nadal, J. R. (2007). Tourism and long-term growth a Spanish perspective. *Annals of tourism research*, *34*(3), 709-726. https://doi.org/10.1016/j.annals.2007.02.003

Pavlyuk, D. (2011). Application of the Spatial Stochastic Frontier model for analysis of a regional tourism sector. *Transport and Telecommunication*, *12*(2), 28–38.

Pérez-Rodríguez, J. V., & Acosta-González, E. (2007). Cost efficiency of the lodging industry in the tourist destination of Gran Canaria (Spain). *Tourism Management*, *28*(4), 993–1005. https://doi.org/10.1016/j.tourman.2006.08.007

Ricardo, D. 1817. On the Principles of Political Economy and Taxation. Variorum edition in Sraffa, P., ed., Works & Correspondence of David Ricardo, Vol. I. *Cambridge: Cambridge University Press,* 1951. https://doi.org/10.1142/9789812701275_0014

Ripoll-Zarraga, A. E., & Raya, J. M. (2020). Tourism indicators and airports' technical efficiency. *Annals of Tourism Research*, *80* (October 2019), 102819. https://doi.org/10.1016/j.annals.2019.102819

Robinson, P. M. (1988). Root-N-Consistent Semiparametric Regression *Econometrica*, *56*(4), 931–954. https://doi.org/10.2307/1912705

Roget, F. M., & Rodríguez González, X. A. (2006). Occupancy level and productivity in rural tourism establishments: The case of Galicia, Spain. *Tourism Economics*, *12*(2), 279–289. https://doi.org/10.5367/000000006777637485

Schmidt, P., & Sickles, R. C. (1984). Production frontiers and panel data. *Journal of Business & Economic Statistics*, *2*(4), 367-374. https://doi.org/10.1080/07350015.1984.10509410

Sellers-Rubio, R., & Casado-Díaz, A. B. (2018). Analyzing hotel efficiency from a regional perspective: The role of environmental determinants. *International Journal of Hospitality Management*, *75*, 75–85. https://doi.org/10.1016/j.ijhm.2018.03.015

Smith, A. (1789). An Inquiry Into the Nature and Causes of the Wealth of Nations. V. 1. A. Straham and T. Cadell.

Solana-Ibáñez, J., Caravaca-Garratón, M., & Para-González, L. (2016a). Two-stage data envelopment analysis of Spanish regions: Efficiency determinants and stability analysis. *Contemporary Economics*, *10* (3), 259–274. https://doi.org/10.5709/ce.1897-9254.214

Solana-Ibáñez, J., Caravaca-Garratón, M., & Para-González, L. (2016b). Stability analysis and efficiency determinants: empirical evidence. *Applied Economics Letters*, *24*(9), 658–661. https://doi.org/10.1080/13504851.2016.1218422

Song, M., & Li, H. (2019). Estimating the efficiency of a sustainable Chinese tourism industry using bootstrap technology rectification. *Technological Forecasting and Social Change*, *143*, 45–54. https://doi.org/10.1016/j.techfore.2019.03.008

Torrens, R. (1815). Essay on the External Corn Trade. London: Hatchard.

Tsionas, E. G., & Kumbhakar, S. C. (2014). Firm heterogeneity, persistent and transient technical inefficiency: A generalized true random-effects model. *Journal of Applied Econometrics*, *29*(1), 110-132. https://doi.org/10.1002/jae.2300

Tzeremes, N. G. 2019. Hotel productivity: A robust Luenberger productivity indicator. *Tourism Economics*, 25(6), 987-996. https://doi.org/10.1177%2F1354816618821271

Tzeremes, N. G. (2020). Robust Malmquist productivity measurement: evidence from Spanish hotel industry during the Great Recession. International Journal of Productivity and Performance Management. https://doi.org/10.1108/IJPPM-01-2019-0037

Wu, J., Cheng, X., & Liao, S. S. (2019). Tourism forecast combination using the stochastic frontier analysis technique. *Tourism Economics*. https://doi.org/10.1177/1354816619868089

Zhou, B., Xu, Y., & Lee, S. K. (2019). Tourism development and regional production efficiency: Evidence from southwestern China. *Tourism Economics*, *25*(5), 800–818. https://doi.org/10.1177/1354816618810250

Tables:

Table 1: Summary of literature review

Authors	Scope	Country	Period	Methodology	DMU	Inputs	Outputs	Determinants
Ben Aissa and Goaied (2016)	Micro	Tunisia		DEA	Hotels	Indirect expenses	Total turnover	Not applied
Chang et al. (2017)	Micro	-		Slacks-based network DEA	Cruise lines	Payroll and related, total operating expenses, marketing, depreciation, non-operating expenses	Passenger ticket revenue, other revenue, operating income net income	Not applied
Deng et al. (2020)	Micro	Macao and Vegas	2008-2016	DEA	Casinos	Employees, total assets	Total expenses, revenue, tax receipts	Financial data, GDP growth, corruption perceived
Huang (2018)	Macro	China	2012	Hybrid network DEA	Regions, Tourist schools, Hotels, Travel agencies	Employees, total assets	Revenues, number of tourists	Not applied
Kularatne et al. (2018)	Micro	Sri Lanka	2010-2014	DEA	Hotels	Employees, number of rooms, assets	Room revenue, other revenues	Age, stars, size dummy, type dummy, Energy, water and waste indices.
Ripoll-Zarraga and Raya (2020)	Micro	Spain	2009-2013	SFA	Airports	Labour cost, operating costs, depreciation	Number of passengers, air traffic movements; cargo, commercial revenues	Not applied
Fernández et al. (2018)	Micro	Spain	2009-2016	SFA	Airports	Capital invested, labour cost, size	Passengers, Cargo, airport revenue	Not applied
Cuccia et al. (2017)	Macro	Italy	2004-2010	Order-m method	Hotels	Accommodation capacity, Tourist arrivals	Tourist bed-nights	Cultural participation indexes
Karakitsiou et al. (2018)	Macro	Greece	2002-2013	DEA	Hotels and restaurants	Number of local units, employees, investments	Turnovers	Not applied
Zhou et al. (2019)	Macro	China	2005-2014	SFA	Regions	Employees, physical capital, tourist arrivals, FDI, technological capital, education expenditure, industry share, public sector share.	GDP	Not applied

Song and Li (2019)	Macro	China	2011-2016	DEA	Regions	Fixed assets investment, employees, tourist spots	Tourist arrivals, tourism revenues	Not applied
Alberca-Oliver et al. (2015)	Micro	Spain	2010	DEA	Trade shows	Net surface area, number of exhibitors	Number of visitors	Not applied
Arbelo-Pérez et al. (2019)	Micro	Spain	2008-2014	SFA	Hotels	Price of labour, Price of materials, prices of other operations, prices of capital	Net sales revenue, other operating revenues	Not applied
Benito et al. (2014)	Macro	Spain	2002-2010	DEA	Hotels	Accommodation capacity, Tourist arrivals	Number of bed-nights	Tourism attractions
Pérez-Rodríguez and Acosta-González (2007)	Micro	Spain	1991-2002	SFA	Hotels	Labour cost, depreciation over total assets ratio, financial expenses over debt ratio	Operating revenue	Not applied
Sellers-Rubio and Casado-Díaz (2018)	Macro	Spain	2008-2016	DEA	NUTS 1 regions	Number of hotels, number of beds, employees	Average daily rate, revenue per room, occupancy rate	Length of stay, international tourists, number of hotels with quality distinctions, sun and beach
Solana-Ibañez et al. (2016a)	Macro	Spain	2005-2013	DEA	NUTS 1 regions	Accommodation capacity, number of beds, tourist arrivals	Overnights, length of stay	Coast, cultural properties, museums, golf clubs, restaurants, retailers
Solana-Ibañez et al. (2016b)	Macro	Spain	2005-2013	DEA	NUTS 1 regions	Number of beds, number of tourists	Overnights	Coast, cultural properties, museums, golf clubs, restaurants, retailers
Deng et al. (2019)	Micro	Spain	2014	SFA	Hotel chains	Average room price, average food price, number of rooms, total assets, material expenses, employee expenses, number of employees, financial expenses, funds, cash flow, operating expenditure, number of establishments	Operating revenue	Share of 3 stars or less in a chain, share of beach hotels, share of golf hotels and share of hotels close to airport

Martínez-Roget and Rodríguez-González (2006)	Micro	Spain	1996-2001	SFA	Rural establishments	Number of workers, number of beds, labour expenditure, cost per room.	Real value added	Not applied
Cordero and Tzeremes (2018)	Micro	Spain	2010-2012	DEA/ Productivity decomposition	Hotels	Employees, fixed assets	Sales	Not applied
Tzeremes (2020)	Micro	Spain	2004-2013	DEA/ Malmquist	Hotels	Employees, fixed assets	Revenues	Not applied
Inchausti-Sintes et al. (2020)	Macro	Spain	2002-2016	SFA/CGE	NUTS 2 regions	Employees, capital stock	Gross Value added	Not applied
Barros (2006)	Micro	Portugal	1998-2002	SFA	Hotels	Salaries, price of capital, market share, operational cost	Sales	Not applied
Chen (2007)	Micro	Taiwan	1995-2001	SFA	Hotels	Total cost, price of labour, price of F&B, price of materials	Total revenue, occupancy rate, value produced per F&B space	Not applied
Cracolici et al. (2008)	Macro	Italia	2001	SFA	Regions	Public cultural patrimony and heritage per population, tourist school graduates over population, share of tourism employees, number of beds	Overnights over population ratio	Not applied
Pavlyuk (2011)	Macro	Baltic states	2005-2008	SFA		Number of beds, employees, museums, area, population, roads per area ratio	Number of tourists	Not applied

Table 2: Showing the main statistics of the variables used in the models

Variable	Obs	Mean	Std. Dev.	Min	Max
Total annual foreigner overnights[1]	800	533,6000	9,311,386	266,519	58,712,364
Average annual hotel beds[1]	800	26,149.76	35,448.68	2,883.167	192,493.2
Average annual labour at hotels[1]	800	3,637.39	5,345.98	281.50	29,601.25
Seasonal concentration[1]	800	0.187	0.100	0.041	0.495
Share of industry[1]	800	0.196	0.073	0.061	0.399
Share of services[1]	800	0.242	0.048	0.158	0.453
Share of exports over GVA[2]	800	0.181	0.143	0.019	0.833
Number of tourist attractions[3]	800	337.88	264.32	64	1592
Share of 4 and 5 star hotels over total accommodation[3]	800	0.057	0.036	0.014	0.207

Sources: 1 INE; 2 DATACOMEX; 3 Tripadvisor

Table 3: Main global results.

	Coefficient	Std. Err.	Coefficient	Std. Err.
	Random Effects model output			
	Model 1		Model 2	
Capital	0.740***	0.039	0.740***	0.039
Capital^2	-0.632***	0.168	-0.632***	0.168
Capital*Time	-0.047***	0.007	-0.047***	0.007
Time	-0.021***	0.003	-0.021***	0.003
Time^2	0.003***	0.000	0.003***	0.000
Financial Crisis	-0.045***	0.009	-0.045***	0.009
Sovereign Debt	-0.074***	0.007	-0.074***	0.007
Sigma U	0.249		0.249	
Sigma E	0.068		0.068	
	Determinants of persistent efficiency variance			
Seasonal Concentration	-5.079***	1.752	-7.656***	1.752
Share of Industry	-7.742***	1.675	-	
Share of Services	-50.456***	5.224	-	
Exports/VAB	-		-0.618	0.649
Tourist Attractions per km^2	-		-72.213***	11.090
4 & 5 star hotel over total supply	-		11.175**	5.100
	Average efficiency coefficients			
Transient Efficiency	0.974	0.000	0.971	0.007
Persistent Efficiency	0.862	0.102	0.849	0.121
Overall Efficiency	0.841	0.104	0.824	0.104

Table 4. Average efficiency scores by cluster

	Efficiency	Avg. Score	Std. Err.
Canary Islands (Las Palmas and Sta. Cruz de Tenerife)	**Transient**	0.972	0.004
	Persistent	0.998	0.000
	Overall	0.970	0.004
Balearic Islands (Baleares)	**Transient**	0.972	0.004
	Persistent	0.999	0.000
	Overall	0.972	0.004
Mediterranean destinations (Alicante, Almería, Girona, Málaga, Sevilla, Tarragona and Valencia)	**Transient**	0.972	0.006
	Persistent	0.965	0.033
	Overall	0.938	0.032
Big Cities (Madrid and Barcelona)	**Transient**	0.972	0.005
	Persistent	0.994	0.006
	Overall	0.966	0.008
Centre Cluster (Álava, Burgos, Castellón, Guadalajara, Guipúzcoa, La Rioja, Navarre, Palencia, Soria, Teruel, Toledo, Valladolid and Zaragoza)	**Transient**	0.971	0.008
	Persistent	0.816	0.096
	Overall	0.793	0.093
‘Others’ (Albacete, Asturias, Ávila, Badajoz, Cáceres, Cádiz, Cantabria, Ciudad Real, Córdoba, Coruña, Cuenca, Granada, Huelva, Huesca, Jaén, León, Llleida, Lugo, Murcia, Ourense, Pontevedra, Salamanca, Segovia, Vizcaya and Zamora)	**Transient**	0.974	0.008
	Persistent	0.803	0.116
	Overall	0.781	0.113

Table 5: Model comparison.

	Coefficient	Std. Err.	Coefficient	Std. Err.	Coefficient	Std. Err.
	Lien, Kumbhakar and Allen (2018)		**Green (2005)**		**Battese and Coelli (1995)**	
Capital	0.740***	0.039	0.741***	0.038	0.580***	0.209
Capital^2	-0.632***	0.168	-0.552***	0.162	-0.094	1.140
Capital*Time	-0.047***	0.007	-0.039***	0.006	-0.053***	0.020
Time	-0.021***	0.003	-0.020***	0.003	-0.027***	0.006
Time^2	0.003***	0.000	0.003***	0.000	0.004***	0.001
Financial Crisis	-0.045***	0.009	-0.044***	0.008	-0.045***	0.009
Sovereign Debt	-0.074***	0.007	-0.064***	0.008	-0.048***	0.013
Sigma U	0.249		0.053		0.197	
Sigma E	0.068		0.055		0.188	
			Determinants of efficiency variance			
Seasonal Concentration	-7.656***	1.752	-24.607***	2.768	-7.036**	2.794
Exports/VAB	-0.618	0.649	1.953	1.289	-0.798	1.406
Tourist Attractions per km^2	-72.213***	11.090	-79.254**	30.802	-65.890***	19.678
4 & 5 star hotels over total supply	11.175**	5.100	-8.703	7.724	-2.370	9.268
			Average efficiency coefficients			
Transient Efficiency	0.971	0.007	-	-	-	-
Persistent Efficiency	0.849	0.121	-	-	-	-
Overall Efficiency	0.824	0.104	0.960	0.047	0.854	0.115

Figures:

Figure 1: Map of share of services of the Spanish NUTS 3 regions.

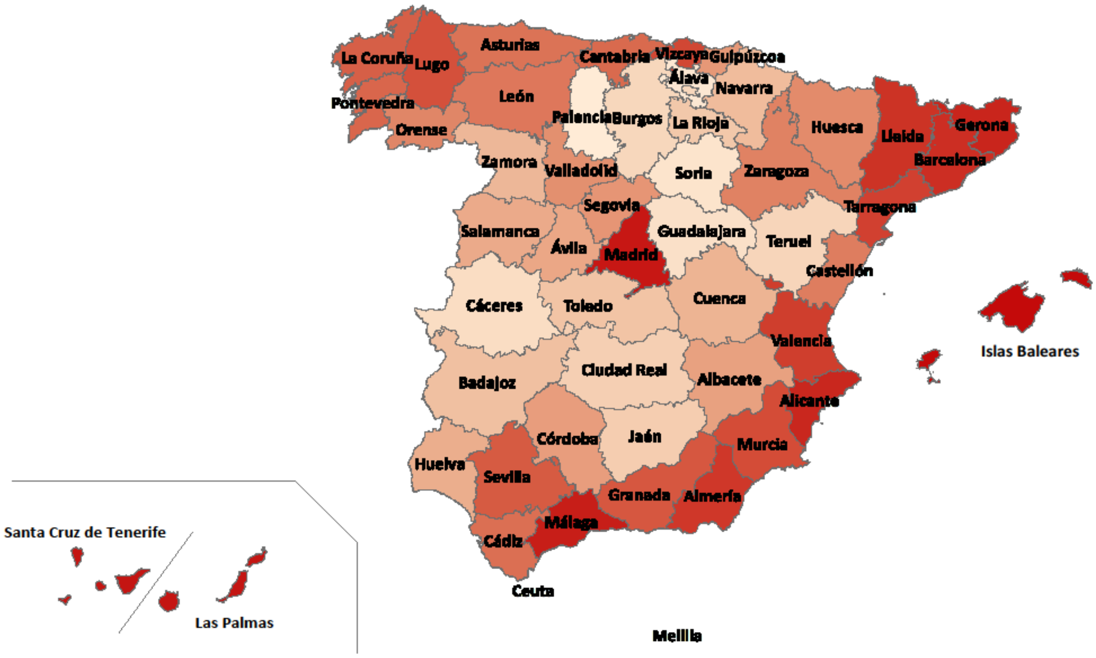

Figure 2: Polar graphic of international overnights

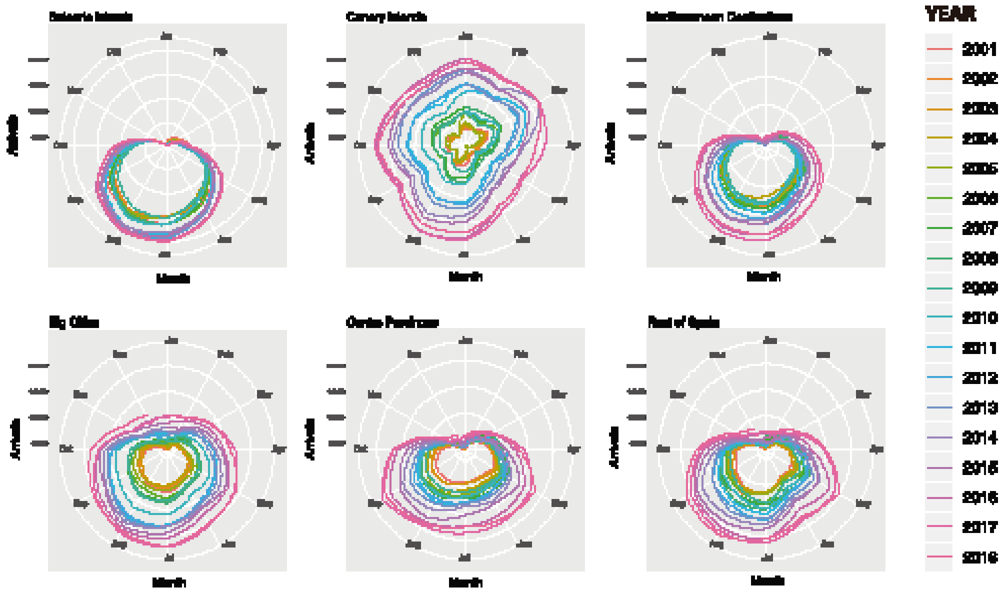

Figure 3: Map of efficiency scores of the Spanish NUTS 3 regions.