Assessing the Ability of Regions to Attract Foreign Tourists:

The Case of Italy

Bernardina Algieri
University of Calabria

Antonio Alvarez
University of Oviedo

Abstract

This study examines the ability of Italian regions to attract foreign tourists and the factors influencing the choice of regional destinations using a novel Stochastic Frontier Demand Model. The results show that several factors including climate, seasonality, cultural heritage and infrastructures influence tourism performance. Easy accessibility to World Heritage Sites drives international tourism demand too. On average, Southern regions lie below the stochastic frontier and are inefficient, while Northern regions tend to be efficient. Calabria, Sardinia and Molise have a low ability to entice foreign tourists, while Veneto maximizes the number of tourist arrivals, overnight stays and expenditures.

Keywords: Regions, Tourism Demand, Stochastic Frontier, Efficiency
JEL Classification: C51; R11; Z32

Introduction

Tourism plays an important role for the development of regions. It contributes to economic growth, creates employment opportunities, and fosters regional progress (e.g., Croes 2014; Eugenio-Martín et al., 2004). Given the importance of this sector, it is relevant to investigate the potential of regions to attract visitors. We define this potential as the maximum number of tourists a region can attract, given its specific characteristics including climate, location, infrastructures, and cultural heritage.

For this purpose, we use a Stochastic Frontier Model in a demand framework. In this setting, the frontier represents the maximum amount of foreign tourist flows a region can entice when its resource endowment and characteristics are efficiently utilized. This approach allows us to distinguish between regions that are on the frontier and those below it. Regions on the frontier are classified as efficient meaning that they are able to attract the maximum number of tourists (or generate the highest amount of revenues), given their resource endowment. Conversely, regions below the frontier are inefficient since they are not fully exploiting their resources and could attract more visitors or increase revenues. It is worth noting that the efficient regions are not necessarily those that attract more tourists in absolute terms, but rather those regions that maximize the number of tourist arrivals given their own endowment: a region that attracts a relatively small number of tourists may be efficient because this small number is the utmost it could expect given its resources.

The present study extends the existing literature on the ability of regions to attract tourists and generate revenues. While previous research has measured this ability in a production framework (e.g., Cracolici et al., 2008), no studies have adopted a Stochastic Frontier Analysis (SFA) within a demand structure. In this setting, we treat visitor and revenue flows as the result of tourism demand by consumers rather than an output produced by regions as the traditional approach does. Frontier demand functions are not common in the applied literature and, to our knowledge, there are only a few studies on energy consumption (Filippini and Orea, 2014; Lin and Wang, 2014; Filippini and Hunt, 2011).

The analysis focuses on Italy, the fifth-largest recipient of tourists in the world. Italy is an interesting case to examine because the country has a long-standing tradition in tourism since the Grand Tour of the 18th Century described by Goethe in his Italian Journey[1]. The country is characterized by a strong North-South divide. Although internal economic disparities are evident in almost every country in the EU, Italy presents particularly high regional contrasts in terms of GDP per capita, unemployment rate, export values and tourism performance (ISTAT, 2021). Despite Italy recorded a tourism gross value added of about 10% of the EU total, this value is lower than those registered by Spain (27%) and Germany (12%) (Eurostat, 2020) and the country shows low competitiveness compared to Greece, Spain and Portugal (Algieri et al., 2018). Thus, improvements in efficiency from a regional perspective are needed to heighten the competitive position of tourism sector (Martín et al., 2017).

We use a quarterly panel dataset of Italian regions for the period 1997-2018, and introduce a modelling novelty by allowing for different effects of seasons across regions. This is important because some regions receive more tourists in summer (sun and beach destinations) while others receive more tourists in winter (well-known winter sport resorts).

Another difference with other studies on tourism demand in Italy is that we just consider foreign tourists. Most studies use domestic tourists (e.g., Massidda and Etzo, 2012), while very few (e.g., Canale et al., 2019) attempt to model the behaviour of foreign visitors at the regional level. The importance of foreign visitors for the tourism industry and the current account of regions and countries has been previously acknowledged (García-Sánchez et al., 2013). Our analysis explicitly focuses on foreign tourist activity measured in terms of arrivals, overnight stays and expenditures. While tourism flows are generally evaluated in terms of arrivals (Patuelli et al. 2013), these three dimensions of tourism activity provide a more comprehensive view of the phenomenon.

[1] Italienische Reise.

In addition to a broad set of regional characteristics, including climate, location, infrastructures and cultural endowments, we introduce a new determinant of tourist flows, namely the degree of accessibility to the World Heritage Sites in each region. Previous studies (De Simone et al., 2019; Cuccia et al. 2017) have considered the number of World Heritage Sites as an explanatory factor of regional tourist demand. We believe that the ease of access to a particular location is an important characteristic that has to be taken into account. The Eolie Islands, in the North of Sicily, included by UNESCO as a World Heritage natural site in 2000, are a good example to understand that their ability to attract tourists is limited by the considerable difficulty to reach them.

The remainder of the study is organized as follows. Section 2 discusses the existing literature on the evaluation of efficiency in tourism. Sections 3 and 4 present the theoretical model and describe the considered data. Sections 5 and 6 illustrate the empirical model and show the econometric estimations. Section 7 discusses the policy implications. Section 8 concludes.

Tourism and regional efficiency

The importance of tourism for economic growth is well documented in the literature (e.g., Menegaki and Tiwari, 2021; Brida et al., 2016; Sinclair, 1998). Many studies have examined this relationship at the country level (e.g., Balaguer et al., 2002) and the regional level (e.g., Cortés-Jiménez, 2008). Previous analyses have not only demonstrated the importance of tourism for regional GDP growth (Paci and Marrocu, 2014), but also for improving the socio-economic conditions of urban and rural areas (Andraz et al., 2016)[2] and reducing regional welfare inequalities (Chaabouni, 2019).

Few studies to date have examined 'tourism efficiency', that is the capacity of micro or macro-units (hotels, regions) to attract tourists given their characteristics. Generally, non-parametric methods based on Data Envelopment Analysis (DEA) and parametric techniques based on Stochastic Frontier Analysis have been used

[2] It must be recognized that some exceptions exist. For example, Figini and Vici (2010) did not find any relationship between tourism specialization and economic growth for a set of countries.

to assess efficiency. Both methods follow the same basic process: first, a frontier is estimated, and then an efficiency index is calculated for each unit as the distance between the actual output and the frontier. We have chosen the SFA approach because it allows the estimation of both the regional efficiency levels and the demand function parameters. Furthermore, SFA has the advantage over the DEA approach in that accounts for measurement errors and other statistical noises.[3]

The majority of the research on tourism efficiency has been performed at the microeconomic level, considering hotels or restaurants (e.g., Dong et al. 2020; Chaabouni, 2019; Detotto et al., 2014; Arbelo et al., 2017). Analyses at the macro-level, focusing on the efficiency of regions in attracting tourists, are, however, relatively scarce. This is rather surprising given the increasing concerns of local authorities about the capacity of regions to become attractive tourist destinations. At this macro-level, Benito et al. (2014) estimated a production frontier using DEA to assess the efficiency of Spanish regions in increasing tourists, and reported substantial differences across areas. Botti et al. (2009) applied the DEA methodology to evaluate the tourism performance of the twenty-two French regions for the year 2006. The authors found that ten out the twenty-two regions are technically efficient.

With reference to the Italian case, Suzuki et al. (2011) used DEA to evaluate the performance of Italian provinces in luring visitors and found that the most efficient destinations are in the North and the central area of the country. The South and the Islands, instead, show efficiency indices less than 0.5. Cuccia et al. (2017) also adopted DEA to examine the role of UNESCO sites in enhancing tourism demand for the Italian regions and uncovered a not significant role of world heritage sites in fostering technical efficiency.

To our knowledge, only two studies use a Stochastic Frontier analysis to analyse the efficiency of Italian regions. Cracolici et al. (2008) adopted a Stochastic Frontier (complemented by a Data Envelopment Analysis) to evaluate the tourist efficiency of Italian destinations. The authors documented that technical efficiency varies greatly across Italian provinces. Artistic and cultural destinations perform better than coastal or

[3] See Assaf and Josiassen (2015) for a review of frontier studies in the tourism literature.

mountainous destinations. Cuccia et al. (2016) explored the effects of cultural heritage in fostering tourism demand in Italian regions. The authors used a DEA approach, accompanied by an SFA as a robustness check. Both methods indicate that cultural heritage sites do not increase tourism demand.

We fill the gap in the existing literature first by examining the ability of Italian regions to attract foreign tourists and generate revenues using an SFA in a novel demand framework not yet used in tourism research. Second, we incorporate some macro-regional variables to account for differences in the proximity to northern border. These macro-regional dummies have not been previously used in the literature. Third, we include innovative contextual drivers of tourism demand, such as the degree of accessibility to the World Heritage Sites in each region. Fourth, we provide a finer analysis using quarterly data to capture the impact of seasons and we allow the seasonal effects to vary across regions by interacting the summer and winter effects with dummies for having coast or winter sports resorts, respectively.

Theoretical model

To estimate the efficiency level of each region, we use the stochastic frontier approach, originally developed by Aigner et al. (1977). This is one of the primary methods used for measuring production efficiency at the firm level. In production theory, a firm is technically efficient when it yields the maximum output that the technology allows for a given set of inputs (Greene, 1993). We apply this production concept to regions, but depart from previous efficiency studies, by framing our stochastic frontier analysis within a demand setting. Thus, a region will be considered efficient if it attracts as many tourists as possible (i.e., it generates the maximum demand) given its characteristics.

In our basic model, consumers decide how to spend their income on different goods and services including travelling for holiday and leisure. Let Q_{ij} be the quantity demanded of tourism (trips or overnight stays) by consumer *j* in destination *i* (country, region, city). The demand function can be written as:

$$Q_{ij} = D_j(P_i, I_j, Z_i) \qquad (1)$$

where P is a measure of prices at the destination, I is the income of consumer *j* and Z is a vector of destination characteristics (e.g. temperature, coast, places to visit).

The aggregate demand for tourism in region *i* is the sum over all individuals that have demand for region *i*:

$$\sum_j Q_{ij} = \sum_j D_j(P_i, I_j, Z_i) \tag{2}$$

Therefore, the aggregate demand function for region *i* can be written as:

$$Q_i = D(P_i, I_i, Z_i) \tag{3}$$

where I_i is some measure of the income of the visitors that have tourism demand for region *i*. The prices of substitutes are not considered to have a more parsimonious specification.

Figure 1 displays the demand functions for two regions. The y-axis reports the quantity demanded of tourism to be consistent with equation (3). D_1 is the demand for region 1, which is associated with a level Z_m^1 of characteristic m (where income and other characteristics are being held constant) and D_2 is the demand for a region with a higher value of characteristic m. For any value of P (conditional on everything else), the quantity demanded of tourism in region 2 is higher than in region 1. Therefore, regional characteristics (as well as income) act as shifters of the demand function. Assuming that the demand is increasing in Z, for given prices and income, regions with a higher value for characteristic Z_m will have a higher tourism demand.

[Figure 1.]

To complete our modelling framework, we move from this theoretical setting to an empirical one. We assume that for a given Z, some regions are inefficient and lie below their frontier. This inefficiency is captured by a non-negative region-specific random term u_i as follows:

$$Q_i = D(P_i, I_i, Z_i) - u_i \tag{4}$$

The frontier mirrors the absence of inefficiency, i.e., u=0, while regions with u>0 are inefficient. The inefficiency of each region is measured with respect to its own frontier. This means that each region is compared with the frontier determined by regions with similar characteristics. For example, regions with coast are compared with the frontier for regions with coast.

Figure 2 displays two demand frontiers D_1 and D_2, which correspond to two different endowments of Z. There are four regions: A_1 and B_1 with a resource endowment Z_1, and C_2 and E_2 with a resource endowment Z_2. Regions A_1 and E_2 are efficient. B_1 and C_2 are instead inefficient regions. C_2 is more inefficient than B_1 as it is further away from its frontier (i.e., $u_C > u_B$), even though C_2 attracts more tourists than B_1 for the same level of prices. E_2 is more efficient than C_2 albeit the region attracts fewer tourists than C_2.

Equation (4) is known as a deterministic frontier function since all deviations from the frontier reflect inefficiency. This strong assumption, however, neglects the random nature of economic variables. Aigner et al. (1977) developed the concept of stochastic frontier to allow for random noise in the dependent variable, by adding a symmetric random term with zero mean, v_i. Therefore, the specification of the stochastic frontier demand function becomes:[4]

$$Q_i = D(P_i, I, Z_i) + v_i - u_i \tag{5}$$

[Figure 2.]

The frontier demand model can also be used in a DEA approach. However, for the empirical implementation of our model we prefer SFA to DEA for the following two reasons. First, SFA allows to account for statistical noise. Since several of the factors that influence tourists' decisions will not be able to be measured, we feel more comfortable with a model that accounts for statistical noise instead of assigning the effect of unmeasured variables to inefficiency. Second, our objective is not only the estimation of efficiency levels for the Italian regions but we are also interested in the interpretation of the estimates of the demand function. We want to know if the included variables are significant or not, as well as their marginal effects. This can only be done in a parametric framework.

Data

[4] The frontier demand specification in the energy studies cited in the Introduction has a composed error term (*v+u*) since inefficient households consume energy above the frontier (i.e., they waste energy). Our error term is (*v-u*) since inefficient regions have tourist demand below their frontier.

For our empirical analysis, we have collected quarterly data for the 20 Italian regions from the National Institute of Statistics (I.STAT), the Bank of Italy, UNESCO and meteo.it. The period of analysis goes from 1997 Q1 to 2018 Q4.

The dependent variable

Tourist activity, our dependent variable, can be measured in different ways. While most studies use a single measure, we consider three proxies, as in Taylor and Ortiz (2009) or Pompili et al. (2019): the number of tourist arrivals[5], the number of overnight stays and tourist expenditure[6].

Table A.1 in the Appendix displays the values for the three variables in the initial and final years of our sample and their percentage change in that period. The regions of Lazio, Lombardy and Veneto are the main tourist destinations during the considered time frame. Conversely, Molise and Basilicata are the less attractive destinations for foreign visitors. Interestingly, the evolution of the three variables differs widely across regions between 1997 and 2018. The regions that register a contraction in the number of foreign tourists are Abruzzi (-9.9%), Calabria (-13.5%), Marche (-31.7%), Molise (-51.7%), Trentino Alto Adige (-20.3%) and Umbria (-7.4%). Eight regions record a reduction in the total overnight stays during the same period, while ten regions register a contraction of tourism expenditure in real terms.

The dynamics of tourist flows, overnight stays and real expenditure for the whole country are reported in Figure 3. Tourist arrivals increased steadily over time, especially after 2005. Overnight stays remained relatively stable until 2010 and then started growing. Real expenditure surged in 2013, just after the sovereign debt crisis was over. It is interesting to note that the number of tourists has increased more than overnight stays and real receipts, which is likely consistent with the general behaviour of consumers during the financial crisis.

[Figure 3.]

[5] Tourist arrivals include overnight and one-day visitors.

[6] Tourist expenditure includes accommodation, restaurants and cafés, transport within Italy, purchases of goods in shops and other services (e.g., museums, concerts and shows, guided tours, vehicle rental).

The explanatory variables

The specification of a demand function requires different types of explanatory variables: prices, income and a set of regional characteristics that control for observed heterogeneity.

a) Prices

- There is no data at the regional level on prices that can represent the different price levels of tourist-related products across regions. This is a common problem in tourism demand studies. The main elements of price for tourists are the cost of travel and the cost of living in the destination. Since these data are not often available, the usual practice is to incorporate the regional Consumer Price Index (CPI) (e.g., Massidda and Etzo, 2012).[7] Thus, we include the CPI at the regional level (2018=100) as an explanatory variable (PRICE).[8]

b) Income

- The income of the tourists is not observed. Our data are aggregated at the regional level and therefore we ignore the country of origin of the tourists. In such setting, it is not possible to account for income. If we could identify tourists by country of origin, the country's per capita GPD could be used to account for differences in income. Other papers that work with data sets similar to ours (Canale et al., 2019) do not include an income variable either. Since most tourists visiting Italy come from the Euro area, we thought of using an average per capita income weighted by each country's share of the arrivals at the border. However, this variable would be common to all regions and could be confounded with the time trend. Hence, as in other studies, we exclude an explicit variable for income and include a time trend and a dummy for the financial-economic crisis. The time trend can partially pick up the evolution of per-capita income over time. The dummy crisis (D_CRISIS), that takes value 1 for the years from 2008 until 2015, would capture the effect of a sharp decrease in per-capita income recorded during the years of the world crisis.

---

[7] When there is a single destination, it becomes easier to find a price variable. For example, Fujii and Mak (1981) used plane fares from New York and San Francisco to explain tourist arrivals in Hawaii.

[8] The real exchange rate is another variable that reflects the cost for tourism, however, this variable has been excluded since the majority of tourists visiting Italy come from the Euro area (over 70% in 2018). Additionally, the estimation including CPI shows a better goodness of fit than the specification with the real exchange rate.

c) Climatic variables

- It is well-known that climate affects the choice of regional destination by tourists (e.g., Eugenio-Martín and Campos-Soria, 2010). While the most common climatic variable is temperature, other variables have been considered in the literature, namely rain, humidity, visibility, the number of sunshine hours or the presence of thunderstorms. Generally, for tourist demand, climate is the most relevant factor in trip-planning stages, whereas weather becomes more important during the trip. For this reason, we include in our model the monthly average temperature (TEMPERATURE) and rainfall (RAINFALL) in each region throughout the whole sample period. In doing so, we are making temperature time-invariant (over the years) since we are interested in an indicator or climate rather than an indicator of weather.

d) Cultural attractions

- The number of UNESCO World Heritage Sites (WHS) is an indicator of the cultural and historical attractiveness of regions. Italy is the country with most sites inscribed in the World Heritage List. The number of WHS has increased from 10 in 1997 to 54 in 2018. Their regional distribution is shown in Figure 4.[9] Therefore, following previous studies (e.g., De Simone et al. 2019), the total number of WHS is considered as a possible driver of tourism demand. However, the raw count of WHS probably does not measure the 'true' attractiveness of each region in terms of cultural heritage. Some sites are not easy to access in the sense that they are not located within large cities or well-connected to transport hubs. To capture the site 'accessibility', we create a variable that is intended to reflect the ease of reaching a particular WHS. This variable takes on value 3 if the WHS is in a place with good access (e.g., Venice, Naples), value 2 if it has to be accessed by car (e.g., Agrigento, Matera), and value 1 if the place is rather isolated (e.g., Eolie Islands). We create a dummy variable for each of the three categories. We exclude the dummy for isolated places and interact the other two (D_ACCESS2

[9] The number of WHS in 2018 (54) is less than the sum of WHS in all the regions in our sample (70). The reason is that some sites are shared by different regions. Two examples are the *Sacri Monti* monuments, which are located in Lombardy and Piedmont, and the site denominated *Longobards in Italy*, which comprises buildings in six regions.

and D_ACCESS3) with the number of WHS in order to make the effect of this variable dependent on the degree of accessibility.

e) Natural features

- Many visitors appreciate the beauty of the seaside and the opportunity to take advantage of beaches. We expect that regions on the coast receive more tourists than inland regions. Following Cuccia et al. (2016), we use the number of kilometers of beaches in each region (BEACHLINE). The kilometers of coastline (Cafiso et al., 2018) produced similar results in the estimation of the frontier but penalized some regions with long coastline, especially the Islands, since many km of coast are not available for tourist use.

f) Infrastructures

Transportation infrastructure is an important facilitator of the movement of people in and between regions and countries (Cafiso et al. 2018; Crouch, 2010). We include a dummy that takes value 1 if the region has an airport and 0 otherwise (D_AIRPORT)[10].

g) Regional dummies

- Regional fixed effects are not included since several explanatory variables are time-invariant. We partly account for individual unobserved heterogeneity by comprising dummy variables for geographical location. D_NORTH, D_CENTER, D_SOUTH, and D_ISLANDS take value 1 for their corresponding regions.[11] These dummies could play an important role. For example, many foreign tourists access the country by car, so that regions that share a border with other countries have a higher probability of receiving foreign tourists. All regions in the dummy North (except for Emilia-Romagna) border another country. The excluded category is D_ISLANDS. To the best of our knowledge, the use of macro-regional variables has not been done in previous papers about the demand for tourism in Italy. Anyway, we must recognize that regional differences have been

[10] We have also estimated a specification with the number of airports, but the results showed a lower goodness of fit.
[11] The North comprises the following regions: Lombardy, Liguria, Piedmont, Emilia Romagna, Friuli-Venezia Giulia, Trentino Alto Adige, Valle d'Aosta and Veneto. The Center includes: Tuscany, Umbria, Marche and Lazio. The South includes: Abruzzi, Molise, Campania, Puglia, Basilicata, Calabria. Finally, Islands comprises Sicily and Sardinia.

considered previously but in different forms. For example, Massidda and Etzo (2012) considered macro-regional differences but splitting the sample into different groups. A typical regional variable in tourism studies with Spanish data is a dummy for the islands (e.g., Priego et al., 2015) but it has not been considered in Italian studies since most of them use bilateral flows tourist data and they include the distance between regions, which already considers the especial location of the islands.

- Following previous studies (e.g., Biagi et al., 2021), we enter a dummy variable equal to 1 for the region that hosts the capital city of the country (D_LAZIO). We add a further dummy for another important region, Lombardy, which contains the second-largest city and airport of the country, Milan (D_LOMBARDY). This is in line with Priego et al. (2015), who included dummies for Madrid and Barcelona in their analysis of regional differences in tourism in Spain.

h) Seasonal dummies

- Tourism demand has a strong seasonal component (Rosselló Nadal et al. 2004) but not very many studies use data at the monthly or quarterly levels. Among those studies which use data at a time frequency lower than years, some do not include seasonal dummies, although they partially account for seasonal variation by including climatic variables which are obviously correlated with seasons (for example, Li et al., 2018). The usual modelling is to include monthly (Muñoz et al., 2021) or quarterly (Maddison, 2001) dummies, although other modelling possibilities exist. Taylor and Ortiz (2009) have monthly data of British regions but they just include a dummy for summertime. We model the effect of seasonality using quarterly dummies in order to control for time effects not explained by other variables. D_QUARTERt (t=1…4) takes on value 1 if the quarter is t, 0 otherwise. Since it is more frequent to travel in summer, we expect that the dummy variable 'D_QUARTER3' to have the largest influence on tourism activity. The excluded category is D_QUARTER4 (October, November and December).
- We further assume that the quarterly effect is not homogeneous across regions. Particularly, we expect the effect of summer to be larger in regions on the coast, while we expect a positive effect of winter in regions with important winter sport resorts. Thereby, we interact the dummy for summer with the dummy for having coastline (D_COASTQ3) and the dummy for winter with a dummy variable

that takes value 1 for the following regions: Trentino, Val d'Aosta, Piedmont and Lombardy (D_WSPORTSQ1).

i) Control variables

- A time trend enters the model to seize the effect of unobserved influential variables which vary over time but are common to all regions, and which have not been captured by other time-varying variables. This is a common variable in most demand studies. Patuelli et al., (2013) justify its inclusion to grasp "long-run change in tourist tastes". To add flexibility to this effect, we include a quadratic term as well as interactions with the regional dummies to capture different trends across regions.

To summarize, our model incorporates prices, climatic variables, cultural attractions, natural features and infrastructures. It also controls for the general trend in tourism activity, seasonality and regional specificities (interaction dummies). A description of the variables used in the empirical analysis is presented in Table A.2 (Appendix).

[Figure 4.]

Empirical Model

The empirical model to be estimated is the following Cobb-Douglas stochastic frontier demand function[12]:

$$lnQ_{\mathrm{it}} = \alpha + \beta \mathrm{lnX_{it}} + \gamma \mathrm{lnW_i} + \sum_{j=1}^{3} \mu_j DQ_{jt} + \theta_1 \mathrm{DQ1}_t * \mathrm{DWSports} + \theta_2 \mathrm{DQ3}_t * \mathrm{DCoast_i} + \delta_t \mathrm{t} + \delta_t \mathrm{t}^2 + \sum_{j=2}^{4} \delta_j DRegion_j * \mathrm{t} + \mathrm{v_{it}} - \mathrm{u_{it}}(Z_{it}) \qquad (6)$$

where *ln* stands for natural log, subscript *i* indicates region and subscript *t* represents time. The dependent variable reflects tourist demand for location *i* in period *t* and is measured as the number of tourists, overnight stays and tourist expenditure. X_{it} is a vector of variables that vary across regions and over-time (WHS), while W_i is a vector of time-invariant regional characteristics (D_AIRPORT, TEMPERATURE, RAINFALL, BEACHLINE)

[12] The Cobb-Douglas, i.e., double-log, functional form is the most common one in demand studies (e.g., Massidda and Etzo, 2012; Pompili et al., 2019) due to allowing for non-linearities and for the ease of interpretation of the estimated parameters as elasticities. Anyway, other functional forms have been used. For example, Maddison (2001) estimated linear and semi-log demand functions.

and some dummies that catch the time-invariant spatial unobserved heterogeneity (D_NORTH, D_CENTER, D_SOUTH, D_LAZIO, D_LOMBARDY). We control for the time dimension of the panel by including a time trend (t), a quadratic trend (t^2) and the interaction of the trend with regional dummies. Finally, seasonality is seized by a set of quarterly dummies (DQ) and we interact the dummy for winter (DQ1) with a dummy for winter sports (D_WSPORTS) and the dummy for summer (DQ3) with the dummy for coast (D_COAST).

The error is composed of two terms: *v* is a symmetric random disturbance which captures the effect of statistical noise and is distributed as a $N(0,\sigma_v^2)$, while *u* is a non-negative random disturbance which captures inefficiency and is assumed to follow a distribution that depends on a set of variables *Z*. There are two alternative specifications of $u_{it}(Z_{it})$, depending on whether *Z* affects the distribution of *u* through its mean or its variance. We have chosen to model the variance of *u* since the presence of heteroskedasticity in *u* will yield biased estimates of both the frontier parameters and the efficiency scores. This result differs markedly from the typical effect of heteroskedasticity in the two-sided error term *v*, which causes the variances of the parameter estimates to be biased.[13] Caudill et al. (1995) incorporate heteroskedasticity into a frontier model, assuming that u_{it} is distributed as $N^+(0, \sigma_{it}^2)$. They assumed that u_{it} exhibits multiplicative heteroskedasticity, a choice that we will use in our analysis.

Variables in the inefficiency term

Our stochastic frontier model allows the assessment of regional inefficiency and the evaluation of the differences in inefficiency levels across regions and over time. The literature does not offer much guidance about how to choose which variables should be included in the *Z* vector. For example, Kumbhakar and Lovell (2000) for the case of production frontiers state that they are "…neither inputs of the production process nor outputs of it, but which nonetheless exert an influence on producer performance". The fundamental idea is to correct efficiency scores for the effect of different environments. That is, the efficiency scores of regions operating in a favourable (difficult) environment should be adjusted downwards (upwards).

The choice of the explanatory variables to include in the inefficiency function is not that easy. We follow two simple rules: (i) to avoid the use of ratios that involve the dependent variables, since that could be a source

[13] Caudill et al. (1995) state that "…the ranking of firms as to their relative inefficiency changes dramatically when the correction for heteroskedasticity is incorporated into the estimation". This is considerable evidence that inefficiency measures are sensitive to heteroskedasticity and must be viewed with caution unless heteroskedasticity is allowed for in the model.

of endogeneity problems; (ii) the variables must be somehow under the control of regional authorities, since one of the basic principles of stochastic frontiers is that the inefficiency is due to the behaviour of the units. With these two rules in mind, the following variables (in natural logs) enter the inefficiency term:

- Accommodation size (HOTELSIZE). It is possible that the size of establishments may affect the efficiency in attracting tourists due, for example, to less search costs compared to small establishments. We try to reflect this aspect by including the average number of bed places per hotel.
- Tourist facilities (FACILITY). Tourists appreciate the existence of a complementary supply of services (e.g., restaurants, shops, sport facilities) that make their stay more comfortable. The presence of these services, which is more likely in large cities and in specialized tourist areas, is a source of economies of agglomeration (Marco-Lájara et al., 2016). The latter could be captured by the number of tourist districts in each region. However, given that this information is not available, we proxy agglomeration considering the number of beds per km2.
- Public safety (HOMICIDES). Other things being equal, more dangerous places are less attractive than safer ones. Previous studies have used violent crime (Marrocu and Paci, 2013) or minor crime (Cafiso et al., 2018) as proxy for public safety. We measure safety by the number of homicides per 100,000 people.
- Congestion (CONGESTION). It has been argued that the high concentration of tourists may act as a deterrent. Previous studies have used population density (Massidda and Etzo, 2012; Marrocu and Paci, 2013) as an indicator of the degree of congestion. We measure tourist saturation in the destination region by the number of beds per thousand inhabitants.
- Time (TREND). There may be other unobserved time-varying factors common to all regions that affect their ability to attract tourists. To catch these latent factors, a time trend is included. 'Learning by tourists' can be one of these factors. For example, while the length of the beachline is fixed, an increase in tourist awareness of the beauties of Italian beaches would result in higher demand, and therefore increase the efficiency of the regions.

Results

The model in equation 6 is estimated by maximum likelihood using Limdep V10. The estimates for each specification (number of tourists, overnight stays and tourist expenditure) are presented in Table 1.

Since the results are similar across the three models, we comment in a general way.

The price coefficient carries a negative sign in the two cases where it is significant. This indicates that price competitiveness is an important factor in tourists' destination choice.

The presence of an airport contributes positively to tourist demand. Airports are seen as a policy lever to boost regions, cities and national economies, and facilitate arrivals (e.g., Blonigen and Cristea, 2015).

Climatic factors are important. Regions with higher temperatures receive more tourists while the opposite happens in the case of rain. These results are consistent with most of the previous studies that include weather variables in tourism demand (e.g., Priego et al., 2015).

Italy's outstanding world heritage sites are a source of revenues and tourist attractors. While tourist demand increases with the number of sites, the accessibility to these locations is important too. The variables with the accessibility-dummy interacted with the number of World Heritage Sites are positive and increasing when accessibility rises (the excluded category is poor accessibility).

As expected, BEACHLINE makes a significant contribution to explaining Italian tourism demand. Sun-and-sea tourism is one of the leading motives in tourists' regional destination choice. The result is consistent with several studies (Benito et al. 2014; Barros et al. 2011).

The macro-regional dummies are positive and significant, indicating that they possess some characteristics not included in the model that make tourist demand higher in those regions than in the Islands (the reference category). The coefficient for the North dummy is the largest, reflecting, among others, the fact that this macro-region must be crossed by visitors travelling by car or bus. The coefficients of the dummies for Centre and South are smaller than that of the North, reflecting the increasing distance to the northern border.

[Table 1.]

The dummies for Lazio and Lombardy, which are positive, would suggest the importance of several cultural touristic cities (Rome, Milan, Bergamo, Mantua, Viterbo) and the fact that they host the two main airports in the country. The variable dummy for Lazio also reflects the fact that the Vatican City is located within Rome and it is a pull factor for many catholic tourists.

The dummy for quarter 1 is negative while the dummies for quarters 2 and 3 are positive and significant. As expected, the coefficient of the dummy for summer (DQ3) is the largest, reflecting that most people travel in July-August-September due to good weather and the usual vacation time. Additionally, the two interactions of the quarterly dummies are significant, indicating that regions that have climate and facilities for winter sports receive more tourists in winter than other regions. Moreover, regions with coast receive more tourists in summer.

The time trend in the frontier could capture, among other factors, the movements in income common to all regions, and changes in consumer tastes. The effect of the trend is not direct since the specification includes not only a linear and a quadratic term but also interaction terms with the macro-region dummies (the Islands is the reference category). The evaluation of the derivative of the dependent variable with respect to the trend in the last year of the sample indicates that in the case of expenditure, the effect of an additional year is positive for the Islands and negative for the other regions. This effect was lower during the years of the economic crisis, as reflected in the negative sign of the dummy variable for the crisis.

Turning to the estimation of the inefficiency term, some interesting results arise. The variable HOTEL SIZE is significant and negative in two specifications. This suggests that having larger hotels enhances the ability of a region to be closer to the demand frontier. This effect can be due to several reasons including a reduction in search costs by tourists or the fact that most large hotels are often multinational chains with high quality, which reduces asymmetric information problems.

As expected, the effect of tourism facilities is always negative and significant, since an increase in this variable indicates more services in the territory.

The homicide rate is not significant in the three models. This result probably indicates that the tourists do not perceive the differences in safety across regions big enough to affect their destination choice.

The variable tourist congestion is always positive and significant, implying that the larger this variable, the further away from the frontier.

Finally, the time trend in the inefficiency term is positive and significant, indicating that some factors common to all regions and different from those included in the model, are causing regional inefficiency to increase over time. The explanation for this finding is not easy, but could be due to some latent factors linked to 'market sentiment' that hinder regional tourist potentials.

Efficiency Analysis

Whereas the estimation of the frontier function only gives an estimate of the composed error term, it is possible to extract the inefficiency component, *u*, using the formula by Jondrow et al. (1982). If the dependent variable is measured in logs, the efficiency (EFF) of region *i* in period *t* is then calculated as:

$$\mathrm{EFF_{it}} = \mathrm{e}^{-\mathrm{u_{it}}} \tag{7}$$

This efficiency index is bounded between zero and 1, with 1 indicating that the region is efficient and is located on its demand frontier. The estimation of the stochastic frontier gives the variances of the error components, σ_u^2 and σ_v^2. Table 1 shows that the variance ratio λ (=σ_u/σ_v) is equal to 0.65 (0.67, 0.70) when the dependent variable is visitors (nights, expenditure). A value of λ less than 1 indicates that the random noise is more important than inefficiency in explaining differences in output across regions.

Table 2 displays the results of the estimated regional efficiency indices EFF (eq. 7) over the sample period. Average efficiency is similar in the three models, ranging from 73.0% in the expenditure model to 74.1% in the overnight stays model. These results mean that, on average, regions could improve their demand by approximately 26% using the same level of resources.

An important point is the interpretation of the efficiency indexes. They are conditional on the explanatory variables included in the frontier part of the model, which means that those variables cannot explain the differences in efficiency levels across regions. Since inefficiency (u_{it}) is part of the error term of the demand function, it captures part of the effect of omitted variables, since v_{it} seizes part of it. Although we have been very careful with the specification of the model, trying to incorporate as many relevant variables as possible, there are always some overlooked aspects, such as differences in quality. For example, two regions may have the same length of coastline but one can have beaches which are more beautiful, resulting in higher attractiveness for tourists.

[Table 2.]

Veneto, Liguria and Campania are among the top 5 most efficient regions in the three models. The score of Veneto is mainly due to the attractiveness of Venice, which goes way beyond being a WHS. Liguria is located very close to two populated French cities (Marseille and Nice), while Campania, known by the Romans as "Campania Felix" (Fertile land), is very famous for its cultural and historical beauties, and its fabulous coasts, Sorrento, Positano, Amalfi, the Gulf of Naples, and the isles of Capri, Ischia and Procida. Friuli, Lombardy and Trentino are also among the top positions in the three models. Veneto is the most efficient region in terms of tourists (98.7%) followed by Friuli (94.4%) and Liguria (92.1%). Veneto is the most efficient region in terms of overnight stays (97%) and tourism receipts (97.7%) too. The excellent performance of Veneto is not surprising as it contains not only Venice, but also Verona. The region has good transport connectivity, and the most impressive art museums in Italy, such as Peggy Guggenheim Collection, the National Art Gallery and Correr Museum. Trentino Alto Adige is situated in the very north of Italy bordering Austria and Switzerland, and is best known for its peaks and mountains: The Dolomites, Brenta and the Gardena Valleys. Hundreds of miles of ski slopes make this region a cutting-edge tourist destination.

The least efficient regions are Sardinia, Molise and Calabria, despite their potentially longer seaside vacation season (except for Molise) than many other regions of the country. In the three models the average levels of efficiency were estimated to be below 60%. These regions share some common features: they are in the

South and located far away from the national border and they have smaller airports than other regions in the country.

Our results document significant levels of inefficiency in the tourism sector for most Southern regions, excluding Campania and Sicily, while Northern regions turn out to be more efficient, except for Valle d'Aosta. The South, hence, does not fully exploit its potential given its resources and there is room for policy interventions to promote the attractiveness of Southern tourist destinations.

There are some possible explanations for the larger inefficiency of the South. The main one is probably connectivity. Southern regions are more difficult to be reached by car, high-speed rail and have smaller airports than the Northern regions. Albeit, the largest airport in the country is located in Rome, it is likely used by tourists who then travel to the Central and Northern regions. For instance, two important tourist locations, Pisa and Florence, are not that far from the Capital. Conversely the high-speed rail system ends up in Salerno and the regions of Calabria, Puglia, Basilicata, and Molise show marked disadvantages in accessibility.

It is important to note that the efficiency index for each region is calculated with respect to the region's frontier. That is, there is not just a common frontier for all regions. On the contrary, the different characteristics of each region define its own frontier. Therefore, the differences in efficiency across regions cannot be due to the different resource endowments but to other factors, such as differences in quality.

An important question is how efficiency evolves over time. Our three models pick up the time-varying efficiency levels between 1997 and 2018. Figure 5 shows that the efficiency patterns in the expenditure and stay models increased from 1998 to 2005, the dynamics for visitors are relatively stable over time.

[Figure 5.]

The temporal patterns of efficiency in the four macro-regions (Figure 6) reveal very similar information: while the efficiencies of the Central and Northern regions are rather stable over time, with a slight decreasing trend, the efficiency of the Southern regions and Islands has increased during the sample period. This holds true especially for Sicily and Sardinia.

[Figure 6.]

Discussion section

The results of the empirical analysis are important especially from a policy point of view. Given that regional governments are interested in any information and guidelines to improve the attractiveness of their regions, the results of our model provide guidance with respect to the following questions:

a) What are the factors that increase the demand for tourism?

b) Which are the factors that hamper regions to exploit their attractiveness potential?

With respect to the first question, there are two groups of demand shifters that could enhance the attractiveness of a given destination and rise tourism demand. The first group, comprising regional dummies, the length of coastlines, climatic factors and time controls, cannot be modified by regional authorities. The second group of shifters, instead, can be influenced to a certain extent, by regional policy makers. This group includes the heritage and natural resources (measured by the number of WHS)[14] and the degree of regional accessibility (measured by the number of airports).

Thus, a first indication is that any effort to increase the number of WHS would pay off. This result is important due to the current debate in the literature about the significance of WHS to entice tourists (see Cellini, 2011; Ribaudo and Figini, 2016). Our results show that not only the number of WHS increases the demand for tourism, but also their accessibility. This leads to a second policy suggestion, namely an improvement in transport connectivity, especially in the Southern regions, would be essential to boost tourist demand.

[14] Regional managers cannot "produce" new heritage resources, but they can promote their resources to be included in the World Heritage List and improve the accessibility to these sites.

With respect to the second question, our model gives some answers in the inefficiency part. Obviously, the way to increase efficiency is to reduce inefficiency. However, this is not so easy because regional authorities should be able to identify the sources of underperformance in order to make a better use of their resources and, hence, improve efficiency. Our analysis would suggest that an increase in tourism facilities, less congestion and larger hotel size could, to a certain extent, reduce inefficiency.

Nonetheless, the efficiency indices derived from an estimated frontier, while useful to quantify the extent of inefficiency, are not that helpful when it comes to improving efficiency. The reason stems from the implicit assumption of the frontier models according to which inefficient regions should behave in the same way as those on the frontier with the best practices. The problem, as pointed out by Alvarez and Arias (2014), is that efficiency measures only inform in terms of 'how much', but not on 'how' regions are actually behaving. For instance, if an inefficient region changes its characteristics (e.g., has one more WHS) but continues operating in the same manner as before (e.g., without defining proper marketing plans), then the region will be able to attract more tourists but its inefficiency will likely to increase (because its frontier is higher).

Our results can be affected by the Covid-19 pandemic. The tourist flows almost disappeared during 2020 although they are recovering in 2021. Anyway, the demand for tourism after Covid may be affected by the higher preference of tourists for security. Therefore, the probability of choosing not only destination but also type of accommodation will be affected by the new preferences by tourists formed during the pandemic (Aiello et al., 2020).

Conclusions

This study has examined the ability of Italian regions to attract foreign tourists and the factors influencing the choice of regional destination for the period 1997-2018. To this purpose, a stochastic frontier model within a novel demand setting has been estimated. Stochastic frontier demand models have been only used in energy studies. This is the first application to tourism demand. In our model, tourism activities have been gauged in terms of foreign tourist arrivals, overnight stays and expenditures at a regional level. Each tourism demand specification controls for price, cultural heritage sites, infrastructures, geographical area, climate, and a time trend. The results show that tourists tend to lean on these variables when they select a specific destination. The number of World Heritage Sites and the ease of access to them are positive and significant in our models. The importance of several renowned cities in the regions of Lazio and Lombardy represents a pull factor for tourism too. Seasonality plays a significant role in driving tourism demand.

Moreover, Veneto, Liguria and Campania are the most efficient regions, while Calabria, Sardinia and Molise are the least efficient regions. On average, Southern regions do not fully exploit their potential given their resource endowments. The results associated to the inefficiency term, reveal that the average size of hotels and the lack of congestion (more comfort services) increase the attractiveness of the regions, while crime is not a significant factor.

From a policy perspective, a number of recommendations can be drawn. Firstly, an easy accessibility to regions is a strong factor that influences tourism demand for Italian destinations. Secondly, factors linked to regional tourism supply such as culture, and transport infrastructure can give an important competitive advantage to a tourism destination. Thirdly, the potentially longer seaside vacation season for the regions of Calabria, Sardinia, Puglia and Sicily should be better exploited.

Improving regional efficiency may not be a simple task due to the lack of complete information about the specific source of underperformance. If, on the one hand, more tourism facilities, less congestions and larger hotel sizes could lessen inefficiency, on the other, a reasonable strategy may be to imitate the successful stories implemented by the most efficient regions. This benchmarking procedure is not without its problems since the strategies adopted in some places cannot fit to other contexts. Therefore, choosing the appropriate reference region becomes an important decision which may have to be subject of specific analytical tools. We consider this to be an important research line for the future.

From a methodological point of view, an interesting extension of the current paper would be to compare the estimation of regional efficiency indexes using the demand and the production frontier approaches. The dependent variable in both approaches is the same but since the explanatory variables would be different, the results could show changes not only on the levels but also on the rankings of the estimated efficiency indexes.

References

Aiello F, Bonanno G, Foglia F (2020) On the choice of accommodation type at the time of Covid-19. Some evidence from the Italian tourism sector. *Current Issues in Tourism*, DOI: 10.1080/13683500.2020.1846504

Aigner D, Lovell CAK, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics*, *6*, 21–37.

Algieri B, Aquino A, Succurro M (2018) International Competitive Advantages In Tourism: An Eclectic View. *Tourism Management Perspectives*, 25, 41–52.

Alvarez A, Arias C (2014) A selection of relevant issues in applied stochastic frontier analysis. *Economics and Business Letters*, 3, 3–11.

Andraz JM, Norte NM, Gonçalves HS (2016) Do tourism spillovers matter in regional economic analysis? An application to Portugal. *Tourism Economics,* 22(5), 939-963.

Arbelo A, Pérez-Gómez P, Arbelo-Pérez M (2017) Cost efficiency and its determinants in the hotel industry. *Tourism Economics* 23(5), 1056–1068.

Assaf AG, Josiassen A (2012) Identifying and Ranking the Determinants of Tourism Performance: A Global Investigation. *Journal of Travel Research*, 51(4), 388-99.

Assaf AG, Josiassen A (2015) Frontier analysis: A state-of-the-art review and meta-analysis. *Journal of Travel Research*, 55(5), 612–627.

Balaguer J, Cantavella-Jorda M (2002) Tourism as a long-run economic growth factor: The Spanish case. *Applied Economics,* 34(7), 877–884.

Barros CP, Botti L, Peypoch N, Robinot E, Solonandrasana B, Assaf A. (2011) Performance of French destinations: Tourism attraction perspectives. *Tourism Management,* 32, 141–146.

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.

Biagi B, Brandano MG, Ortega-Argiles R (2021) Smart specialisation and tourism: Understanding the priority choices in EU regions, *Socio-Economic Planning Sciences*, 74, 100883.

Blonigen BA, Cristea AD (2015) Airports and urban growth: evidence from a quasi-natural policy experiment. *Journal of Urban Economics*, 86, 128–146.

Botti L, Peypoch N, Robinot E, Solonadrasana B. (2009) Tourism destination competitiveness: the French regions case. *European Journal of Tourism Research*, 2(1), 5-24.

Brida JG, Cortes-Jimenez I, Pulina M (2016) Has the tourism-led growth hypothesis been validated? A literature review, *Current Issues in Tourism*, 19(5), 394-430.

Cafiso G, Cellini R, Cuccia T (2018) Do economic crises lead tourists to closer destinations? Italy at the time of the Great Recession. *Papers in Regional Science*, 97, 369–386.

Canale RR, De Simone E, Di Maio A, Parenti B (2019) UNESCO World Heritage sites and tourism attractiveness: The case of Italian provinces. *Land Use Policy*, 85, 114-120.

Caudill SB, Ford JM, Gropper DM (1995) Frontier Estimation and Firm-Specific Inefficiency Measures in the Presence of Heteroskedasticity. *Journal of Business and Economic Statistics*, 13, 105–111.

Cellini R (2011) Is UNESCO recognition effective in fostering tourism? A comment on Yang, Lin and Han. *Tourism Management,* 32(2), 452–454.

Chaabouni S (2019) China's regional tourism efficiency: A two-stage double bootstrap data envelopment analysis. *Journal of Destination Marketing & Management*, 11, 183-191.

Cortés-Jiménez I (2008) Which Type of Tourism Matters to the Regional Economic Growth? The Cases of Spain and Italy. *Int. J. of Tourism Research*, 10, 127–139.

Cracolici MF, Nijkamp P, Rietveld P (2008) Assessment of Tourism Competitiveness by Analysing Destination Efficiency. *Tourism Economics*, *14(2)*, 325–342.

Croes R (2014) The role of tourism in poverty reduction: An empirical assessment. *Tourism Economics* 20(2): 207–226.

Crouch GI (2010) Destination Competitiveness: An Analysis of Determinant Attributes. *Journal of Travel Research,* 50(1), 27–45.

Cuccia T, Guccio C, Rizzo I (2016) The effects of UNESCO World Heritage List inscription on tourism destinations performance in Italian regions. *Economic Modelling*, 53, 494–508.

Cuccia T, Guccio C, Rizzo I (2017) UNESCO site and performance trend of Italian regional tourism destinations: a two stage DEA window analysis with spatial interaction.*Tourism Economics*, *23(2)*, 316–342.

Detotto C, Pulina M, Brida JG (2014) Assessing the productivity of the Italian hospitality sector: A post-WDEA pooled-truncated and spatial analysis. *Journal of Productivity Analysis, 42*, 103–121.

De Simone E, Canale RR, Di Maio A (2019) Do UNESCO World Heritage Sites Influence International Tourist Arrivals? Evidence from Italian Provincial Data. *Social Indicator Research, 146*, 345–359.

Dong H, Peypoch N, Zhang L (2020) Do contextual factors matter? Evidence from Chinese hotel productivity with heterogeneity. *Tourism Economics*, 26(2), 257-275.

Eugenio-Martin JL, Campos-Soria JA (2010) Climate in the Region of Origin and Destination Choice in Outbound Tourism Demand. *Tourism Management*, 31 (6), 744-53.

Eugenio-Martin JL, Martín Morales N and Scarpa R (2004) Tourism and Economic Growth in Latin American Countries: A Panel Data Approach. FEEM Working Paper No 26.

Eurostat (2020) Tourism Satellite Accounts (TSA) in EuropeFigini, P, Vici, L (2010) Tourism and growth in a cross section of countries. *Tourism Economics* 16(4): 789–805.

Filippini M, Hunt L (2011) Energy Demand and Energy Efficiency in the OECD Countries: A Stochastic Demand Frontier Approach. *The Energy Journal,* 32(2), 59–80.

Filippini M, Orea L (2014) Applications of the stochastic frontier approach in Energy Economics. *Economics and Business Letters*, 3(1), 35–42.

Fujii ET, Mak J (1981) Forecasting tourism demand; Some methodological issues. *Annals of Regional Science,* 15, 72–82.

García-Sánchez A, Fernández-Rubio E, Collado MD (2013) Daily Expenses of Foreign Tourists, Length of Stay and Activities: Evidence from Spain. *Tourism Economics,* 19(3): 613-630.

Greene WH (1993), The econometric approach to efficiency analysis, in Fried et al eds, The Measurement of Productive Efficiency: Techniques and Applications, Oxford University-Press, NY, 68–119.

Istat (2021) *Indicatori territoriali per le politiche di sviluppo* Rome.

Jondrow J, Lovell CAK, Materov I, Schmidt P (1982) On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics,* 19, 233–238.

Kumbhakar S, Lovell CAK (2000) *Stochastic Frontier Analysis* Cambridge.

Li H, Goh C, Hung K, Chen JL (2018) Relative climate index and its effect on seasonal tourism demand. *Journal of Travel Research*, 57(2), 178-192.

Lin B, Wang X (2014) Exploring energy efficiency in China's iron and steel industry: a stochastic frontier approach. *Energy Policy,* 72, 87–96.

Maddison D (2001) In Search of Warmer Climates? The Impact of Climate Change on Flows of British Tourists. *Climatic Change*, 49(1), 193-208.

Marrocu E, Paci, R (2013) Different tourists to different destinations Evidence from spatial interaction models. *Tourism Management,* 39(12), 71–83.

Martín JC, Mendoza C, Román C (2017) Regional Spanish Tourism Competitiveness A DEA-MONITUR approach. *Region, European Regional Science Association*, 4(3), 153–173.

Massida C, Etzo I (2012) The determinants of Italian domestic tourism: A panel data analysis. *Tourism Management*, 33(3), 603–610.

Menegaki AN, Tiwari AK (2021) The stability of interaction channels between tourism and financial development in 10 top tourism destinations: Evidence from a Fourier Toda-Yamamoto estimator. *Tourism Economics*, doi:101177/13548166211021174

Muñoz C, Alvarez A, Baños J, (2021) Modelling the Effect of Weather on Tourism: Does it Vary across Seasons? *Tourism Geographies*, 23, 1-22.

Nunes P, Cai M, Ferrise R, Moriondo M, Bindi M (2013) An econometric analysis of climate change impacts on tourism flows: an empirical evidence from the region of Tuscany, Italy. *International Journal of Ecological Economics & Statistics*, 31(4), 1-20.

Paci R, Marrocu E (2014) Tourism and regional growth in Europe. *Papers in Regional Science,* 93(1), 25–50.

Patuelli R, Mussoni M, Candela G (2013) The effects of World Heritage sites on domestic tourism: A spatial interaction model for Italy. *Journal of Geographic Systems,* 15, 396–402.

Pompili T, Pisati M, Lorenzini E (2019) Determinants of international tourist choices in Italian provinces: A joint demand–supply approach with spatial effects. *Papers in Regional Science*, 98(6), 2251-2273.

Priego J, Rosselló Nadal J, Santana M (2015) The impact of climate change on domestic tourism: a gravity model for Spain. *Regional Environmental Change*, 15, 291–300.

Ribaudo G, Figini P (2017)The puzzle of tourism demand at destinations hosting UNESCO World Heritage Sites: An analysis of tourism flows for Italy. *Journal of Travel Research,* 56(4), 521-542.

Rosselló Nadal J, Riera Font A, Sansó Rosselló A (2004) The economic determinants of seasonal patterns. *Annals of Tourism Research*, 31(3), 697-711.

Sinclair MT (1998) Tourism and economic development: A survey. *The Journal of Development Studies, 34(5)*, 1–51.

Suzuki S, Nijkamp P, Rietveld P (2011) Regional efficiency improvement by means of data envelopment analysis through Euclidean distance minimization including fixed input factors: an application to tourist regions in Italy. *Papers in Regional Science*, 90(1), 67–89.

Taylor T, Ortiz RA (2009) Impacts of climate change on domestic tourism in the UK: a panel data estimation. *Tourism Economics*, 15(4), 803-812.

Appendix

Table A. 1. **Foreign Tourism indicators, absolute values and percentage change (1997-2018)**

	TOURISTS (thousand)			OVERNIGHT STAYS (thousand)			REAL EXPENDITURE (million euro)		
	1997	2018	%Δ	1997	2018	%Δ	1997	2018	%Δ
Abruzzi	438	395	-9.9%	5902	2940	-50.2%	474	160	-66.2%
Basilicata	76	154	103.4%	900	596	-33.7%	51	32	-36.8%
Calabria	353	305	-13.5%	6140	3166	-48.4%	233	189	-18.9%
Campania	1443	4446	208.2%	11946	20206	69.1%	1221	2258	84.9%
Emilia	2532	5266	108.0%	15671	21096	34.6%	2036	1932	-5.1%
Friuli	9273	13633	47.0%	11712	12376	5.7%	1993	1321	-33.7%
Lazio	6362	15155	138.2%	38120	62857	64.9%	6475	7258	12.1%

Liguria	7744	9252	19.5%	12414	16588	33.6%	1154	2227	93.1%
Lombardy	13714	23296	69.9%	33875	51237	51.3%	5322	6585	23.7%
Marche	1010	690	-31.7%	4131	3552	-14.0%	451	236	-47.8%
Molise	52	25	-51.7%	702	244	-65.2%	57	13	-76.4%
Piedmont	2638	6241	136.6%	11371	18424	62.0%	1047	1688	61.3%
Puglia	1196	1869	56.2%	13019	12259	-5.8%	807	613	-24.0%
Sardinia	315	1518	381.1%	5739	10456	82.2%	406	958	135.9%
Sicily	985	4599	366.8%	8979	23397	160.6%	664	1902	186.2%
Trentin	6101	4861	-20.3%	32635	18350	-43.8%	2884	1713	-40.6%
Tuscany	5891	9228	56.6%	32508	39988	23.0%	3745	4398	17.5%
Umbria	519	481	-7.4%	4707	2501	-46.9%	374	177	-52.6%
Valle d'Aosta	891	1088	22.1%	1175	2816	139.6%	133	352	163.9%
Veneto	9005	14437	60.3%	47261	58278	23.3%	5450	5992	9.9%
Italy	**73269**	**119879**	**63.6%**	**306954**	**387331**	**26.2%**	**26260**	**40497**	**54.2%**

Source: Elaborations on Bank of Italy data.

Table A. 2. Description of variables used in the empirical analysis

Variables	**Description**	**Source**
Visitors	Number of foreign tourists (000)	Bank of Italy[+]
Stays	Overnight stays by foreign tourists (000)	Bank of Italy
Real Expenditure	Nominal tourist expenditure in million euros CPI adjusted	Bank of Italy
Price	Regional Consumer Price Index (2018=100)	ISTAT
Temperature	Average temperature in degree Celsius	Meteo.it
Rainfall	Mean precipitations in mm	Meteo.it
WHS	Number of UNESCO World Heritage Sites[++]	Unesco
Beach_km	Number of km of beaches	ISTAT
Airports	Dummy equal to 1 if the region has airport	ISTAT

Note: [+]Data on foreign tourists, extracted from the Bank of Italy, have some advantages over the data provided by the Italian National Statistical Institute (ISTAT). The Bank of Italy data come from surveys at the point of departure and includes all types of accommodations, while the data from ISTAT come from a survey undertaken at registered accommodation establishments and, therefore, leaves out tourists residing in other types of accommodation such as private homes and one-day visitors. Moreover, it does not collect data on travellers' expenditure.
[++]There are two types of WHS: cultural (archaeological sites, castles, churches, historical centres of towns) and natural (forests, mountains). Italy has 50 cultural sites and 5 natural sites in 2021.

Tables

Table 1. Estimation of the frontier demand function

	Tourists		Overnight stays		Expenditure	
	Coefficient	Sd. error	Coefficient	Sd. error	Coefficient	Sd. error
Constant	-5.26655	3.50080	24.8594***	2.61009	16.3978***	3.04591
L_CPI	1.25919	.81945	-3.2070***	.60017	-2.0973***	.70508
L_WHS	.01951***	.00480	.00718*	.00386	.01947***	.00456
L_WHSACC2	.19104***	.03791	.10856***	.02866	.16734***	.03237
L_WHSACC3	.60449***	.03096	.66496***	.02236	.75288***	.02668
L_TEMP	1.98973***	.33430	.88829***	.20379	.84708***	.23758
L_RAIN	-.69365***	.09997	-2.1208***	.08638	-1.77831***	.08960
D_AIRPORT	1.75159***	.09292	1.41758***	.06103	1.69868***	.07406
L_BEACHLINE	.04413***	.00452	.02486***	.00350	.03907***	.00378
D_NORTH	2.52264***	.18561	2.09874***	.12850	2.10790***	.15845
D_CENTER	1.11426***	.18289	1.35557***	.12603	1.20661***	.15657
D_SOUTH	.46404***	.17450	.56920***	.11963	.58442***	.15671
D_LAZIO	.92364***	.06941	.55164***	.04379	.98417***	.04795
D_LOMBARDY	.64285***	.09517	.14113**	.05815	.38651***	.06355
D_Q1	-.31479***	.03170	-.19964***	.02250	-.31601***	.02634
D_Q2	.40316***	.03068	.45208***	.02378	.41614***	.02595
D_Q3	.62959***	.05537	.72989***	.04260	.65022***	.04907
D_WINTSPORT_Q1	.58856***	.05810	.50722***	.04806	.70425***	.04867
D_COAST_Q3	.17132***	.06211	.27460***	.04674	.21862***	.05376
TIME TREND	.00089	.02505	.11836***	.01769	.05984***	.02210
TIME TREND_SQ	.00053	.00041	-.00184***	.00029	-.00057*	.00034
D_NORTH_T	-.02983***	.01015	-.03657***	.00678	-.03598***	.00887
D_CENTER_T	-.03592***	.01065	-.04129***	.00688	-.04664***	.00907
D_SOUTH_T	-.03599***	.01047	-.04628***	.00722	-.05837***	.00936
D_CRISIS	-.09907***	.03417	-.05724**	.02601	-.07277**	.02861
Parameters in variance of u						
Constant symm	-1.9094***	.06629	-2.70383***	.07141	-2.4724***	.06859
Constant one side	-8.49221***	1.80109	4.91860***	1.05868	-.58491	1.20573
L_HOTELSIZE	1.09115***	.36921	-1.72661***	.23550	-.53334**	.25683
L_FACILITY	-2.0869***	.40920	-3.04254***	.27604	-2.6701***	.32793
L_HOMICIDES	-.00141	.04712	.01717	.04615	-.00210	.04319
L_CONGESTION	1.24646***	.29504	2.12251***	.21575	1.68065***	.22752
TIME TREND	.16127***	.02840	.03300**	.01597	.07350***	.01940
$\sigma^2=\sigma_v^2+\sigma_u^2$	0.6499		0.6661		0.7011	

$\lambda=\sigma_u/\sigma_v$	0.6491	0.8491	0.8283

Note: This table reports the estimated parameters and standard errors for eq. 6. Time period: 1997Q1-2018Q4. Obs.1760. *p<0.1 **p<0.01 ***p<0.01

Table 2. Estimated regional efficiency indices for the three dependent variables

Rank Region	Dep. variable Visitors	Rank Region	Dep. variable Overnight stays	Rank Region	Dep. variable Expenditure
Veneto	0.987	Veneto	0.970	Veneto	0.977
Friuli	0.944	Campania	0.933	Liguria	0.920
Liguria	0.921	Liguria	0.929	Campania	0.917
Trentino	0.910	Lombardy	0.919	Lazio	0.903
Lazio	0.909	Lazio	0.918	Friuli	0.893
Campania	0.907	Tuscany	0.866	Lombardy	0.888
Tuscany	0.885	Friuli	0.853	Trentino	0.873
Umbria	0.881	Trentino	0.844	Tuscany	0.869
Lombardy	0.873	Marche	0.833	Emilia	0.803
Piedmont	0.818	Emilia	0.832	Umbria	0.788
Sicily	0.802	Piedmont	0.815	Marche	0.786
Emilia	0.798	Sicily	0.813	Piedmont	0.784
Marche	0.751	Puglia	0.800	Sicily	0.778
Abruzzi	0.744	Umbria	0.727	Abruzzi	0.688
Puglia	0.710	Abruzzi	0.721	Puglia	0.680
Basilicata	0.626	Calabria	0.608	Basilicata	0.512
Valle d'Aosta	0.601	Basilicata	0.492	Valle d'Aosta	0.485
Molise	0.581	Valle d'Aosta	0.423	Calabria	0.484
Sardinia	0.409	Molise	0.411	Molise	0.447
Calabria	0.358	Sardinia	0.404	Sardinia	0.405

Note: This table reports the estimated efficiency indices for eq. 6. Time period: 1997Q1-2018Q4.

Figures

Figure 1. Frontier demand functions

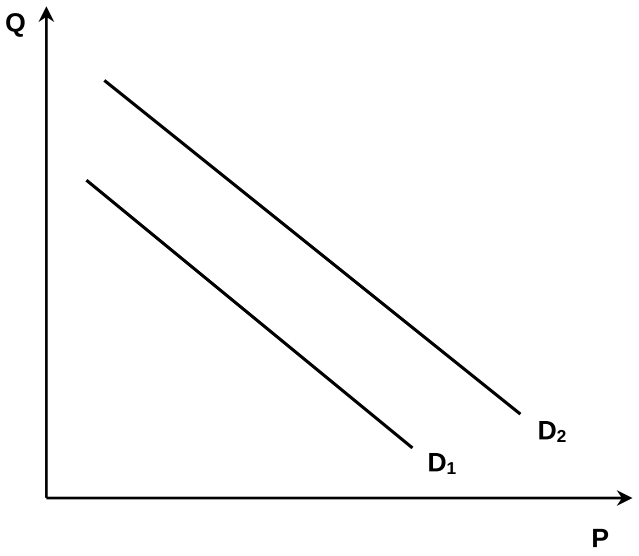


Figure 2. Inefficiency in frontier demand functions

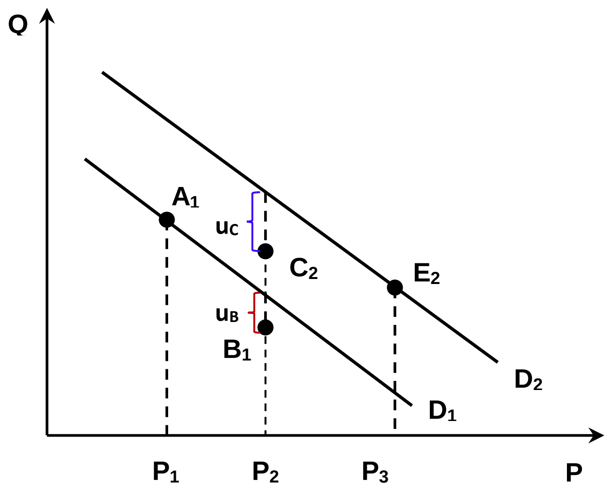
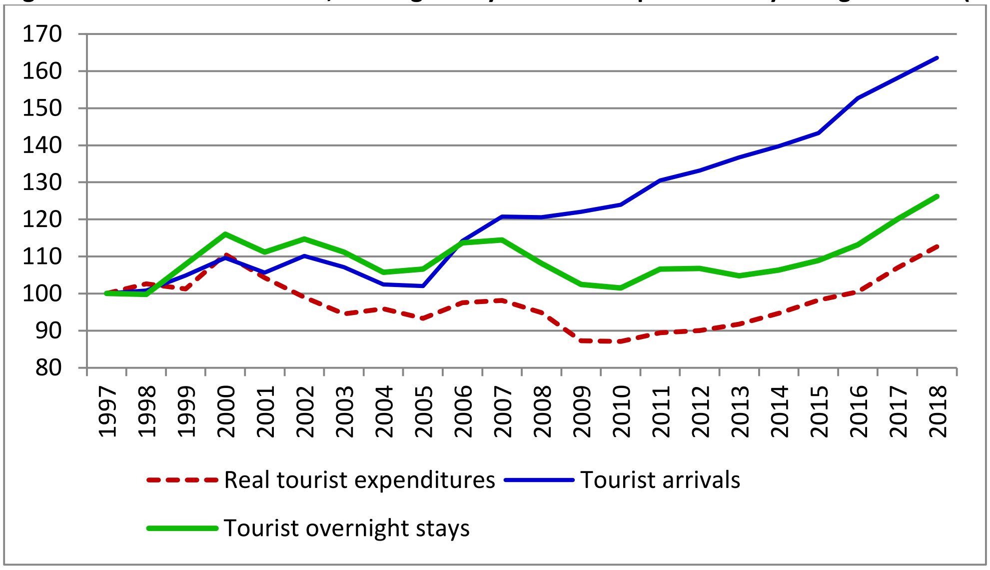


Figure 3. Evolution of arrivals, overnight stays and real expenditure by foreign tourists (1997=100)



Source: Elaborations on Bank of Italy data.

Figure 4. Regional distribution of World Heritage Sites in Italy (2019)

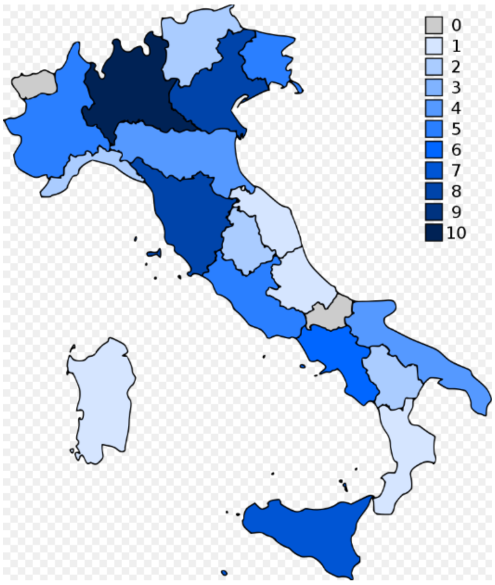


Figure 5. Evolution of the efficiency indices (means over all regions)

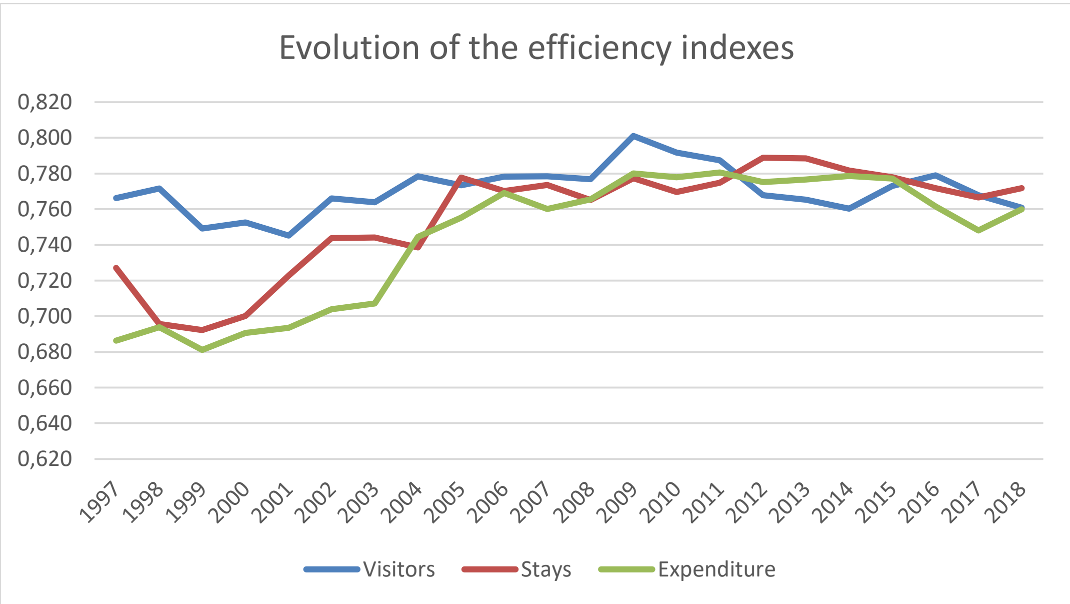
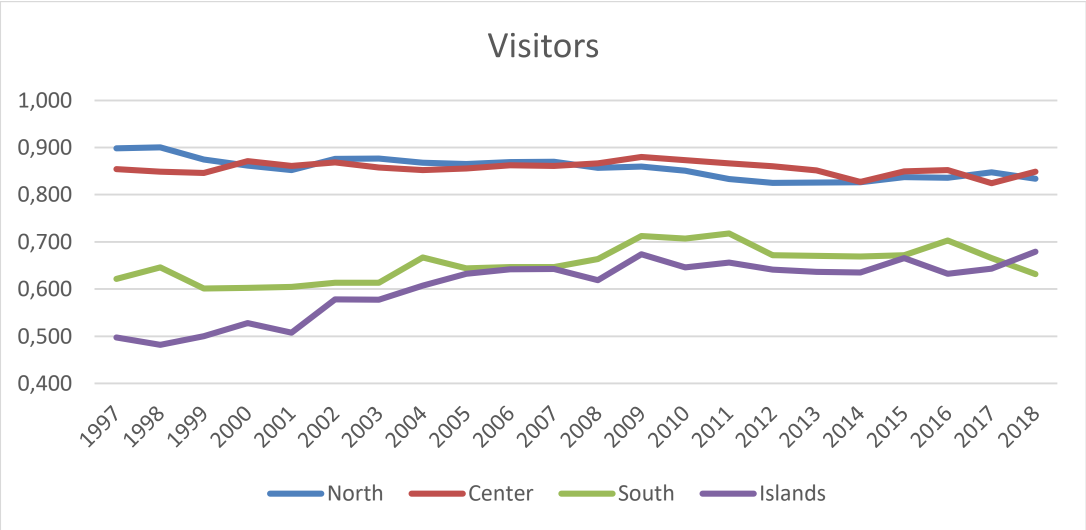
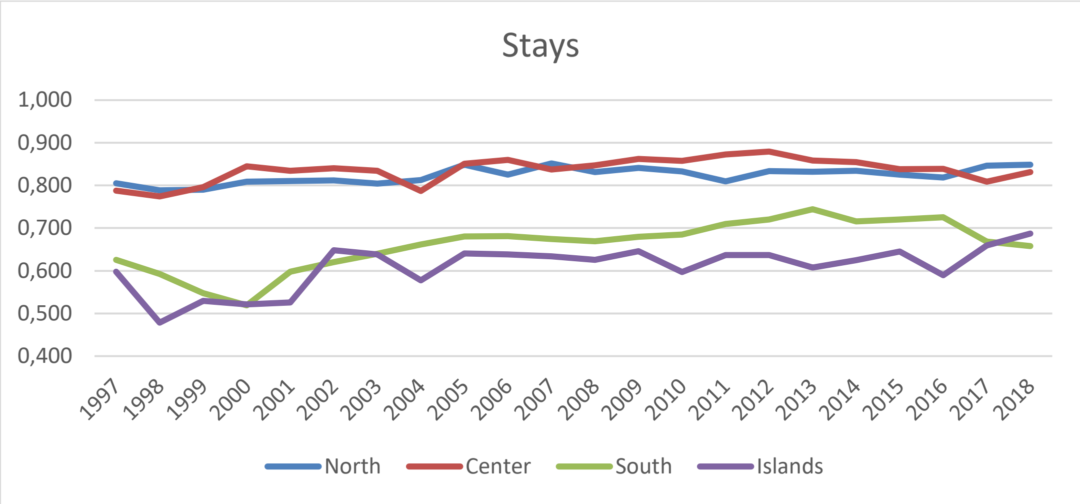
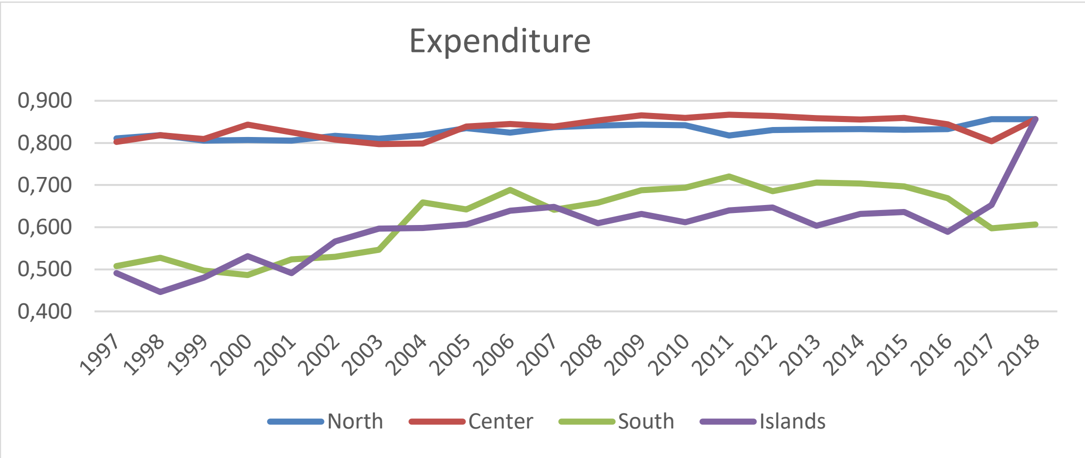


Figure 6. Evolution of the efficiency indices by macro-region




Stays
1,000
0,900
0,800
0,700
0,600
0,500
0,400
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
North
Center
South
Islands


Expenditure
0,900
0,800
0,700
0,600
0,500
0,400
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
North
Center
South
Islands