



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Modelling international monthly tourism demand at the micro destination level with climate indicators and web-traffic data

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Modelling international monthly tourism demand at the micro destination level with climate indicators and web-traffic data / Emili S.; Figini P.; Guizzardi A.. - In: TOURISM ECONOMICS. - ISSN 1354-8166. - ELETTRONICO. - 26:7(2020), pp. 1129-1151. [10.1177/1354816619867804]

Availability:

This version is available at: <https://hdl.handle.net/11585/734949> since: 2022-04-28

Published:

DOI: <http://doi.org/10.1177/1354816619867804>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

Emili, S., Figini, P., & Guizzardi, A. (2020). Modelling international monthly tourism demand at the micro destination level with climate indicators and web-traffic data. *Tourism Economics*, 26(7), 1129-1151.

The final published version is available online at:

<https://doi.org/10.1177/1354816619867804>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>)

When citing, please refer to the published version.

Introduction

Fragmentation of the tourism product (Haugland *et al.*, 2011) and heterogeneity of local suppliers (Sainaghi & Baggio, 2014) are two relevant characteristics of the tourism industry, especially in countries like Italy where the diversification of tourism experiences is very high, as the local communities and their territories (physically and culturally) are usually an essential part of the visitor experience (Goffi & Cucculelli, 2018; Presenza *et al.*, 2013; OECD, 2011). This fragmentation mirrors a strong dependence of the tourism industry on territorial-specific natural and cultural characteristics, attractions and products (Sainaghi & Mauri, 2018), often leading to fierce competition between “small” destinations (Buhalis, 2000). Thus, local destination management organizations (DMOs) - which are in charge of coordination of promotional activities as well as the maintenance and management of cultural and natural attractions (Andergassen *et al.*, 2013; Andergassen *et al.*, 2017) - often call for prompt and spatially detailed statistical information needed to monitor the dynamics of tourism demand (Bornhorst *et al.*, 2010; Sainaghi, 2006).

Unfortunately, the official statistical systems hardly produce tourism information that can be exploited for policy making purposes. The current directive for tourism statistics (UNWTO, 2010) recommends the collection of monthly data on inbound and outbound tourism, at regional (NUTS2) detail. More punctual information (i.e. at provincial or municipal level), is published annually – monthly only in some countries –with a time delay that in the in the more recent Italian official survey ranges from 11 to 23 months (ISTAT, 2018). This “information gap” also regards tourism determinants like income and price levels: GDP is available for NUTS2 areas at quarterly frequency, while price indices are available as a prompt monthly index

number whose representativeness regarding tourists' purchasing power, however, is often questioned (Dwyer & Forsyth, 2011).

Many scholars have recently worked on the idea to bridge this gap by using Big Data (BD), retrievable from the web almost in real time and with the granularity level needed by local policy makers (Mariani *et al.*, 2018; Smith, 2016; Song & Liu, 2017). Specifically, web search queries and climate statistics have been exploited to augment the micro-founded econometric specifications of tourism demand models that embody prices and income as main explanatory variables (Zhang & Kulendran, 2017; Li, Pan *et al.* 2017). The main goal of these and other recent papers is the enhancement of forecasting accuracy, with noticeable results (Pan *et al.* 2012). However, we expect that the inclusion of non-economic data retrieved from the web might also affect the estimate of price and income elasticities, as well as the dynamics of the demand function. The assessment of whether these augmented models are robust in terms of economic interpretation could thereby increase their appeal among scholars and practitioners.

The present paper exactly focuses on this aspect. Accordingly, we aim to shed light on the effects of timely and territorially detailed information (mainly coming from Google Trends (GT), and Composite Climate Indices (CCI)) on tourism demand models' dynamic features, lag dependence structure and elasticities to price and income. We acknowledge that is hard to derive a general conclusion, as many factors influence the econometric relationship between tourism demand and exogenous variables (Peng *et al.* 2015). However, in this paper we control for some of the factors at play by modelling monthly tourism demand (measured in terms of arrivals or overnight stays) in the Italian cities of Catania, Florence and Milan (micro-destinations, representative respectively of the leisure, cultural and business tourism segments). We also distinguish

between two origin markets (Germany and the UK) that differ in terms of currency used. Moreover, we test different hypotheses about dynamics (in terms of seasonality and trend-cycle patterns) and lag/dependence structure. That way, local decision makers are provided with robust empirical evidence about the possibility to exploit web based and climate information to estimate the responsiveness of tourism demand, and to better understand how sensitive different segments are to policy intervention.

Within such a framework, we pose three research questions: are elasticities to price and income robust to the inclusion in the model of complementary variables related to Internet searches and climate conditions? Are dynamics and lag/dependence structures of the augmented models robust across destinations, origin markets and measures of tourism flows? Is there a 'best' Google indicator to be included in the estimation or do tourists in each specific origin/destination combination leave a different fingerprint on the web?

To achieve our research goals, we structure the paper as follows. Section 2 reviews the recent literature on the estimation of tourism demand and the use of big data. Section 3 and Section 4 respectively introduce the model specification and the data. Results are presented in Section 5 while some concluding remarks with useful suggestions for local policy makers are sketched in the last section.

Literature review

The literature on tourism demand rarely considered high territorial granularity and high-frequency temporal data (Song *et al.*, 2009), although tourism is a complex phenomenon where seasonality and specific territorial characteristics (i.e. local attractions) have a dominant role as demand determinants (Smeral, 2014; Gunter & Onder, 2015; Vu & Turner, 2006). Among others, Smeral (2017) lamented the unavailability of exogenous information at the desired disaggregated level as one of the main limitations for advancement of quantitative analysis in the tourism field. Since the 1970s the **modelling of tourism demand** has been characterized by time series analysis dominated by seasonally integrated autoregressive moving-average models (SARIMA), sometimes with the inclusion of GARCH effects or long memory features (Song & Li, 2008, Gil-Alana, 2005).

However, studies where time series models are augmented with economic determinants (mainly income and relative prices) are also **popular**. In this context, findings generally state that tourism is a luxury good (Crouch, 1994; **Munoz *et al.*, 2007; Peng *et al.*, 2015**; Smeral, 2017). Long-haul tourism **displays** a relatively higher elasticity to income because of the more exotic and unique features and the lack of available substitutes (Shiff & Becken, 2011; Peng *et al.*, 2015). However, income elasticity less than one is also present in the literature and might be explained by some “necessary” short-haul international trips (Fuleky *et al.*, 2014). Income elasticity may also vary over time in line with changes in the macroeconomic environment and/or with structural changes in consumer behaviour, leading Smeral (2017) to suggest that tourism could no longer be considered a luxury good. On the contrary, negative income elasticities for "inferior" destinations are rarely reported (Crouch, 1996).

With regards to price, as expected, own price elasticity is often found to be negative, but the magnitude varies considerably depending on the type of tourism and

the time span under consideration (Song *et al.*, 2010). Peng *et al.* (2015) pointed out that in destinations with less substitutes, price competition tends to be less intense, revolving in a lower sensitivity to price. Theoretical exceptions to the law of demand, leading to a positive price elasticity, are possible if the change of a good's price has such a strong impact on purchasing power that it causes a radical change in people's whole pattern of consumption (Crouch, 1992). Sensitivity of demand to price has increased over time due to the major reductions in transport costs (air fares) and to increased competition between destinations (Crouch, 1994) although some market segments, such as business tourism, are less sensitive. Moreover, a prolonged period of low inflation, like the one experienced by European countries in recent years, is likely to boost price elasticity. It is also found that tourists tend to be more aware of exchange rate changes before they travel rather than inflationary effects in the destination they plan to visit (Peng *et al.*, 2015). As researchers can easily access exchange rate data, this variable is sometimes recognized as the best proxy for price dynamics in tourism demand models (Song & Li, 2008).

The impact of economic variables on tourism demand are moderated by other factors. Song *et al.* (2010) and Martins *et al.* (2017) found that tourism arrivals are mainly influenced by income, while tourism expenditure and overnight stays are more affected by the real exchange rate. Crouch (1996) highlighted that estimated elasticities increase with data frequency. Moreover, the lag length of the dependent variable is likely to be associated to a larger absolute value of own-price elasticity (Peng *et al.*, 2015) while no significant difference appears on estimated income elasticities. They also show that the inclusion/omission of other explanatory variables (and the way they are measured) significantly affects the result.

The measurement bias in income and price is often recalled as a limit to the possibility of setting up a reliable inter-temporal relationship with the dependent variable (Song & Li, 2008). Moreover, the potential interdependence between income and price generates a bias in the estimates of income and price elasticities (Seetaram *et al.*, 2016; Peng *et al.*, 2015). From an econometric perspective, both measurement error and collinearity between variables generate an omitted variable bias that is theoretically and empirically recognized as a source of non-spherical models' residuals (Lim 1997) and low forecasting accuracy (Athanasopoulos *et al.*, 2010). This bias increases with the spatial and temporal detail of the analysis, as scholars are forced to proxy local dynamics with national data (Gunter & Onder, 2015). The picture looks even worse when possible measurement errors in the dependent variable (arrivals or overnights) are considered (Guizzardi & Bernini, 2012).

In order to tackle the problems mentioned above, researchers have attempted to introduce non-economic variables in economic models. Ettredge *et al.* (2005) were the first authors to use Google index to analyse the dynamics of the unemployment rate. In the field of tourism, big data are used to improve knowledge about (potential) demand and tourism businesses' target markets (Song & Liu, 2017). However, a large part of the literature uses web-based information to improve forecasting accuracy. Choi & Varian (2012) in their **seminal** paper improved the forecasting accuracy of ARIMA models using travel-related Google search data. Pan *et al.* (2012) obtained the same results with multivariate ARMA, while Bangwayo-Skeete & Skeete (2015) implemented an AR-MIDAS regression. Rivera (2016) used Google Trends data in a dynamic linear model to forecast arrivals in Puerto Rico, while Gunter & Onder (2016) were among the few authors focussing on micro areas (the city of Vienna).

Another strand of literature focuses on the most effective ways to include big data in the econometric model to limit issues of overfitting and multicollinearity. We recognize three main statistical approaches (Li *et al.*, 2017a): the principal component analysis (Li *et al.*, 2015), data shift and summation of different types of search query data, paying attention to the indicators' lag orders (Yang *et al.*, 2015), and the generalized dynamic factor models (Li *et al.*, 2017b).

Finally, a different source of information largely considered by researchers is represented by climatic factors and weather perception (Jeuring, 2017). Both indicators are included either as singular factors (e.g. Falk, 2013) or as composite indicators. A noticeable example is the Tourism Climatic Index (TCI) developed by Mieczkowski (1985) that considers temperature, humidity, rainfall, hours of sunshine and wind speed. Goh (2012) introduced the TCI index to analyse international tourism demand in Hong Kong, reporting a stronger impact of climate conditions when the distance between origin and destination countries increases. Similarly, Li *et al.* (2017a) defined a relative climate index, based on the comparison of TCI between the destination and the origin countries. Zhang & Kulendran (2017) introduced a Composite Climate Index (CCI) where climatic variables in the destination and in substitute countries are weighted by the impact and the volatility of each component.

To the best of our knowledge, climate and search queries data have never been considered together to increase the consistency of income and price elasticities estimation in baseline tourism demand models.

The model(s) specification

We start from the following general specification of the demand model:

$$D = f(\text{time series components, income, prices, climatic indicators, search queries}) \quad (1)$$

where D is the log-difference in tourism demand between same months in subsequent years and $f(\cdot)$ is a linear function. In line with the literature we look at income and prices as the main economic determinants (Crouch, 1992). With regards to income, we consider origin countries' per capita GDP deflated by the corresponding consumer price index (CPI) (eventually multiplied by the exchange rate). We also try to proxy the effect of tourists' income by considering the industrial production index (available at monthly frequency) but, consistently with Song *et al.* (2010) and Dogru *et al.* (2017) we obtain less significant estimates and a worst fit.

As a proxy for price competitiveness, we use the ratio between CPIs in the destination and in the origin countries. To improve the fitness, CPIs at the regional level are preferred to the national CPI. We discard the CPIs for the accommodation sector as they provide less significant estimates and a worse fit. Price levels for substitute destinations are not included in the model because of the theoretical and practical difficulties in defining competing destinations for “micro” destinations (Dogru *et al.*, 2017).

The estimation of eq.1 is repeated for different market segments, seasonality patterns and time-dependence structures, as follows.

Different demands (tourism products)

It is well known (Witt & Witt, 1995; Crouch, 1996) that elasticities are strongly influenced by market segments. Thus, we disentangle the analysis by considering three

destinations and two origin markets. The three cities are representative of different mix of tourism: Florence, a worldwide known art city and cultural destination is mainly an attraction for cultural tourists (Melotti, 2018); Milan, another art city which is also the Italian business and fashion capital (Sainaghi et al., 2018); Catania, a southern destination offering the typical mix between art and Sea & Sun that characterizes most part of Southern Italy (Cuccia and Rizzo, 2011). Demand from Germany and the UK are studied: they were chosen as they are among the top five incoming markets for all three destinations; moreover, they adopt different currencies (Euro and British Pound) allowing to test a possible “exchange rate effect”. We consider both monthly overnight stays and arrivals as alternative dependent variables. We cannot include tourism expenditure into the analysis due to the unavailability of this information at such fine grain territorial level (NUTS3).

The non-economic determinants: climate indices and search queries

Information about climate in the destination is included as a Composite Climatic Index (CCI) following Zhang & Kulendran (2017):

$$CCI_i = \sum_{l=1}^3 w_l \frac{1}{\sigma_l} c_l \tag{2}$$

where w_l is the correlation of each climate indicator c_l , $l = \{rainfall, humidity, temperature\}$ with tourism demand for destination i , while σ_l is the average absolute deviation in relation to the average growth rate of the climatic variable c_l . The CCI is expected to be positively correlated with tourism demand.

Search queries data are obtained from Google Trends as in Bangwayo-Skeete & Skeete (2015). Specifically, searches were defined using Google Trends filters “Category”, “Region”, “Time Interval” and “Web Search” (see the guide at

<https://support.google.com/>). The relative volume of searches containing the name of the three cities (i.e. “Catania”, “Florence”, “Milan”) in the Category “Travels” (and in the sub-categories “*Hotels & Accommodation*”, “*Beaches & Sea*” and “*Historical Sites & Buildings*”) were considered (hereafter: Travel, Hotel, Beaches and Art). Queries were filtered by each origin country (looking specifically at German and UK online searches through the filter “Region”). Collected monthly data include searches returned on all the Google’s search-tools (e.g. Google Images, Google News, etc., defined by the filter “Web Searches”). We expect this indicator to be positively correlated with tourism demand.

Different endogenous and seasonal dynamics

In line with the most common approach in the literature (see Athanasopulos *et al.*, 2010), we model the residual temporal dynamics – i.e. after seasonal differencing – through two autoregressive components at lags 1 and 12. As literature does not offer a clear evidence that seasonality is always stochastic (Coshall, 2005; Sainaghi & Baggio, 2017), we also introduce a deterministic component given by a set of monthly dummies.

Different lag structures for the exogenous variables

For both dependent variables (arrivals and overnights, each measured for the 6 combinations of origins/destinations studied) we estimate baseline models (i.e. without GT and CCI) including income and price variables at the same lag (0, 1, 2 or 12 months), and augmented models which include the CCI (at lag 0, 1, 2, 3 or 12 months) and one among the four GT indices (at lag 0, 1 or 2 months) We lead the specification process by the Akaike Information Criterion (AIC).

The wide set of alternatives described above are represented in the following

generalization of eq. 1.

$$\Delta_{12} \ln D_{i,h,j,t} = \beta_0 + \sum_{m=1}^{12} \beta_m d_m + \varphi_1 \Delta_{12} \ln D_{i,h,j,t-1} + \varphi_2 \Delta_{12} \ln D_{i,h,j,t-12} + \gamma_1 \Delta_{12} \ln(Y |^j / P_j)_{t-k_1} + \delta_1 \Delta_{12} \ln(P |^i / P_j)_{t-k_1} + \theta_1 \Delta_{12} CCI_{i,t-k_2} + \theta_2 \Delta_{12} \ln GT_{i,j,t-k_3} + \epsilon_{i,h,j,t} \quad (3)$$

where $h = \{\text{arrivals, overnights}\}$; $i = \{\text{Florence, Milan, Catania}\}$; $j = \{\text{Germany, UK}\}$; t is a time index and k_n the lag as specified before. Δ_{12} stands for twelfth differences, d_m , with $m = 1, \dots, 12$ are seasonal dummies; Y represents income; P the price index.

We estimate augmented and baseline (i.e. with $\theta_1 = \theta_2 = 0$) models. In both cases eq. 3 is estimated under the following constraints for the endogenous and seasonal dynamics:

- $\beta_0 = 0$;
- $\beta_0 = 0; \varphi_2 = 0$;
- $\beta_m = 0, \forall m$ (4)
- $\beta_m = 0, \forall m; \varphi_1 = 0$;
- $\beta_m = 0, \forall m; \varphi_2 = 0$;

The data

Monthly data on arrivals and overnight stays for each destination disaggregated by origin market for the period January 2004 - December 2014 were provided by ISTAT, the Italian Statistics Office, on May 2017, under fee payment. Monthly demand is characterized by positive trend and seasonal dynamics for each combination of origin/destination under consideration (see Figure 1). British tourists show a single summer peak while Germans display two peaks in May (around Pentecost holidays) and September. Milan has the largest number of arrivals, Florence the largest number of overnights, while Catania has much smaller tourism flows.

[Insert Figure 1 about here]

[Insert Table 1 about here]

Other descriptive statistics are reported in Table 1. Catania records the largest coefficient of variation, indicating a much stronger seasonality than Milan and Florence: the ratio between high and low monthly peaks is 1283% and 815% for German arrivals and overnights respectively, and 623% and 583% for UK arrivals and overnights. The same ratios are much lower in Milan (126% and 133% for Germans, 135% and 145% for Britons), coherently with its business vocation (Sainaghi, Mauri & d'Angella 2018). Florence lies in between.

We tested normality with Doornik-Hansen test, while seasonal and non-seasonal non-stationarity was analysed through Hylleberg, Engel, Granger and Yoo (HEGY) statistics, as in Gunter & Onder (2015). Monthly tourism flows display at least one unit roots at frequencies: $0, \pm \pi/6, \pm \pi/3, \pm \pi/2, \pm 2\pi/3, \pm 5\pi/6$ and π . However, the series in seasonal differences are stationary. None of the dependent variables is normally distributed, with the exception of British demand in Milan.

Information on income and price levels was retrieved from Datastream. Monthly GDP was computed as the third part of the correspondent quarterly figure. Income was deflated by the consumer price index (CPI) of the related source market (P_j). The relative (own) price variable to be included in the model was computed as the ratio between the CPI in the destination region i over the CPI in the origin country j :

$$P_{i,j} = P_i / (P_j / EXR_{i,j})$$

where the exchange rate EXR applied only for British demand.

Climate data were obtained from the Operative Centre for Meteorology of the Italian Aeronautical Military Force. Hourly data for each destination were averaged along each month. Monthly rainfall (mm), temperature (C°) and humidity (%) were combined according to equation (2). As the climatic variables are composite indicators weighed by correlation indices, their summary statistics are not straightforward interpretable.

Google Trends monthly data were available from January 2004: data were summarised as a (relative) index, ranging from 0 to 100, where the upper limit corresponded to the month when the largest number of queries was recorded. Thus, the coefficient of variations (CV) reported in Table 1 can be linked to the product seasonality. In fact, a large CV (hence a large variance and/or a low mean) implies that there are either many months in which the indicator is very low – i.e. the interest of web-people is low – and a few spikes. The existence of GT time series with some zero values (corresponding to months where web search are – relatively – low) prevents to calculate statistics on unit roots. For the remaining indices, the HEGY test results show non-stationarity at least at one frequency.

Results

For each combination of origin/destination ($2 \times 3 = 6$) and for both indices of demand (arrivals and overnight stays) we estimated and compared 24 baseline models (6 models – see eq. 4 – with economic variables specified at lags 0, 1, 2, or 12). Then, for each of the 12 combinations of origin/destination and demand measure we estimated 1440 augmented specifications: the 24 baseline models augmented by considering five different lags for CCI (at 0, 1, 2, 3 or 12 months) and 3 different lags for each of the four GTs (at 0, 1 or 2 months). OLS estimation was performed through a Matlab routine which is available on request.

The large number of estimated parameters suggests presenting the distribution of price and income elasticities in the baseline and augmented models by always setting the non-significant coefficients equal to zero. Detailed estimates are provided only for the 12 best baseline and the 12 best augmented specifications, considering a penalized fit criterion (AIC). The best augmented models highlight some dynamic features of tourists' decisional process for different combinations of origin/destination and arrivals/overnights.

Following Song *et al.* (2010), our benchmark is: for income elasticity the range from 0.54 to 5.4 for German tourists and from 0.48 to 6.02 for UK tourists; for price elasticity the range from -7.4 to -0.18 for German tourists and from -9.9 to 0.95 for UK tourists. In what follows, results are presented for the three destinations separately and then some general conclusions are outlined.

Catania

Catania is a cultural city on the seaside and, considering its whole province, it can be considered as a “*Sea & Sun*” destination, cheaper (in terms of both cost of living and

tourism prices) than Florence and Milan.

[Insert Figure 2 about here]

[Insert Table 2 about here]

Figure 2 (left-hand side – lhs) shows the distribution of the elasticities in the 24 baseline models. Income (measured on the x -axis) has a positive coefficient, which is almost always statistically significant. Elasticities are within the range 2-4 for Germans and 2-5 for UK tourists, indicating that tourists perceive a holiday in Catania as a kind of luxury product: as Catania is not one of the most notorious destinations in Italy, inbound tourists are likely to travel there when they can afford an “extra” holiday. As in Song *et. al.* (2010) we find that arrivals are more sensitive to income than overnights.

Both source markets analysed in this study are wealthy countries, where international travel is a common activity and thus tourists are expected to be less sensitive to price changes. Accordingly, the estimated price elasticities are very low in absolute figures (Figure 2, lhs, y -axis) compared to the values reported in the literature. Findings seem to confirm that tourists are more aware of exchange rate changes than inflation, as we find that price elasticities are significant only for UK arrivals.

The comparison between the baseline (left) and the augmented (center) models highlights that spatially and timely detailed big data do not substantially alter the sign and the significance of the parameters estimated in the baseline models. There are only a few specifications (e.g. UK overnight stays) where positive and significant estimates of price elasticity appear. These specifications, however, have the worst statistical fit.

The best fit has been found when the Beach GT is considered: the own coefficient is almost always significant, and is higher for UK tourists (Figure 2, right-hand side – rhs, y -axis). The CCI index has a significant and positive coefficient only when arrivals are modelled: the straightforward interpretation is that climate is

important for the decision to undertake the trip, not for the decision about the length of stay. Comparing Catania with the other destinations, we observe the lowest coefficients for GT and the highest for CCI, consistently with the idea that climate and weather expectations drive consumers' choice in this "*Sea & Sun*" destination.

The comparison between the best baseline and augmented models (reported in Table 2) shows that the introduction of non-economic variables always improves the overall fit and diminishes the omitted variable bias, as signalled by better residual statistics. The dynamic structure envisages an AR(12) component with a negative sign, indicating that the annual adjustment of the demand to the long-term trend is "oscillating". In the case of the German market the positive AR(1) component represents a cyclical (infra-annual) correction in the return to equilibrium that could be read as the attitude of Germans to take holiday decisions in the short term. Accordingly, the economic variables in the model for Germany enters with the 0 lag, while in the case of the UK, the optimal lag is 2, suggesting that the decision to buy the holiday takes place around two months before the departure.

"Beaches" is the more informative GT index for both markets confirming the Catania's Sea & Sun vocation. Its coefficient is always significant and positive, with very short optimal lags. The best models for Germany include an optimal lag equal to 0 for GT, implying that Google is mostly used to search information on the destination just before the departure. Similarly, the optimal lag in the models for UK arrivals is 1, indicating that this GT index can be considered an important short-term leading indicator for both arrivals and overnights.

The coefficients of the CCI are positive but significant only when demand is measured through arrivals. Consistently with Zhang & Kulendran (2017), the CCI most correlated to the dynamics of demand has a short lag ($k_2 = 2$). This result does not have

a straightforward interpretation, as it would suggest that tourists do not infer weather conditions from climatological statistics of the previous year (same month). In the case of British tourists, the selected k_2 coincides with the lag at which the “average tourist” is buying the holiday in Catania. This suggests that decision makers could be sensitive to news about climatological events when they book the holiday.

Florence

Florence is a worldwide cultural superstar. It is the capital of Tuscany, a region renowned for its countryside, medieval villages and Etruscan sites. Figure 3(lhs) shows that income (*x-axis*) almost always has a positive and significant coefficient in the 24 baseline specifications. Elasticities ranged from 1 to 2.5 for Germany and from 0.9 to 3.5 for the UK. Their values are lower than the corresponding benchmarks in the literature; due to the lack of similar substitutes, the holiday in Florence seems perceived as a “must be” from wealthy countries like Germany and UK. However, while German arrivals and overnight stays have similar elasticities, British arrivals are much more inelastic than overnights. Arguably, income affects more the length of British holidays than the decision to travel in such a unique destination.

[Insert Figure 3 about here]

[Insert Table 3 about here]

Price elasticities (*y-axis*) are consistent with the law of demand for the UK market only; their low absolute values, compared to benchmark, support the idea that Florence is perceived as a “necessary” destination, which visit is worth regardless the price. Concerning the German market, some specifications return positive price elasticities, a result not new for Italy (Konovalova & Vidishcheva, 2013) that we believe is driven by the inadequacy of CPI in representing the relative price of a

holiday. This hypothesis is reinforced by the results obtained for the UK market (negative price elasticity), where the exchange rate is likely to increase the effectiveness in measuring relative prices between countries. A positive elasticity to price can also be connected to the Veblen effect, as reported in Crouch (1992) and the possibility that the income effect might outweigh the substitution effect cannot be excluded as a further explanation.

The comparison between baseline models (lhs) and augmented models (centre) signals that coefficients in the latter are distributed around their respective coefficients in the baseline models, although the estimated elasticities are slightly lower in absolute value, coherently with the increase in the number of regressors.

GT indices are important to explain demand dynamics also in Florence. In particular the optimal GT is “Travel” for Germans and “Art” for the UK (Figure 3, rhs, y-axis). Coefficients are always significant and larger for Germans, especially when demand is measured by overnight stays. The CCI variable is significant only when arrivals are considered, supporting the idea that the climate is not a strong determinant factor for a cultural holiday.

The introduction of non-economic variables improves the overall fit of the models (see Table 3), as the AICs always decrease moving to the augmented specification. Autocorrelation and heteroscedasticity of the residuals cast doubts about the reliability of elasticity estimates and confirm the difficulties of CPI in summarising the role of the relative cost of living for German tourists. On the contrary, the UK models (where the exchange rate brings additional information) show much better residuals statistics.

The dynamic structure presents an AR(12) component with a negative sign, indicating that the annual adjustment of the demand to the long-term trend is counter

cyclical. However, in the case of British tourists, the short-term adjustment to equilibrium is also driven by an infra-annual cycle, probably in association with a greater importance given to a “short-term” Word-of-Mouth effect. Turning to the economic variables, both indices of demand are better explained when a lag of 2 months is considered, and the inclusion of web-based variables does not modify the optimal lag.

The best lag for the GT indicators is always 0, supporting the hypothesis that tourists leave significant fingerprints on the web in the proximity of departure. The relationship among web searches and tourism demand is particularly evident for Germans arrivals. The fact that “Travel” is the category associated to the better specification confirms that the web is used for a general retrieve of information. On the contrary, British seem more aware about things to do and see in Florence, as the sub-indices “Art” and “Hotel” are associated to the best specifications.

The CCI is only significant in modelling the arrivals, not the length of stay. The optimal lag is 12 months for Germany and 3 months for the UK. Therefore, German seem rational, as they form their expectations based on the climate one year before, while British show an intermediate lag (three months, similar to the case of Catania), that is difficult to explain, without assuming that tourists are more influenced by current news about climatological events when they book the journey than by climate statistics.

Milan

Milan is the economic capital of Italy, mainly a business and fashion destination. With direct flights to all the major cities in Europe, it is also less than 4-hour drive from Southern Germany, for which it is certainly a short-haul destination.

[Insert Figure 4 about here]

[Insert Table 4 about here]

Accordingly, Figure 4 (baseline model, lhs) shows insignificant elasticity to income (*x-axis*). The estimated price elasticities (*y-axis*) are only significant for the UK market, where the exchange rate variations allow a better description of the relative cost of living, but their range (from -0.3 to -0.7) confirms the low sensitivity of demand to price that is typical of business tourism.

When spatially and timely detailed big data are included in the augmented models (Figure 4, centre), price elasticities for Germany sometimes become positive and significant. Dealing with business tourism, this evidence may relate to the increased profitability of doing business in a destination, which pushes prices up (Blake & Cortes-Jiménez 2007). However, as both baseline and augmented specifications present residuals' autocorrelation (see Table 4), we believe that the significant and positive price elasticities do not exclude that tourists may adjust their purchasing behaviour upon arrival at the destination to account for unexpected price levels (Seetaram *et. al.* 2016).

The comparison between the best models (Table 4) confirms that the introduction of GT and CCI improves the fit (the AIC decreases). Residuals statistics for the German models highlight a misspecification problem, which is much less serious in the UK case. Accordingly, UK price elasticities are negative and highly significant, confirming that the exchange rate is a better proxy for tourism price dynamics, especially in regime of low inflation. The optimal lag signals that the decision to visit Milan is expected to be taken with a short lag, of about two-month advance.

The dynamic structure presents an AR(12) component with a negative sign, indicating countercyclical adjustment to the long-term trend. In the case of German tourism, the positive AR(1) component represents a month-to-month cyclical adjustment that could be associated with the need for repeat visits during, for example, the execution of a contract, something that is typical for a business destination.

The GT associated to the best fit is the sub-category Travel. Coefficients are almost always significant and larger than in other destinations, especially for the UK segment (Figure 4, rhs). The optimal lag is 0, indicating that tourists are more interested to search for Milan in the proximity of the departure. The CCI variable is rarely significant and the coefficients are generally lower than in Catania and Florence. GT indices are hence the only informative variable for policy makers willing to monitor demand dynamics in a framework of business tourism which is, as expected, not affected by climate conditions.

Conclusions

The present work aims at assessing if and when spatial and timely detailed non-economic big data such as Google Trends data (GT) and Composite Climate Indicators (CCI) embody useful information about tourism demand dynamics in small areas (local destinations), where accurate and timely official statistics are seldom available. As previous literature employs climate and search query variables separately, focusing on the enhancement of forecasting accuracy, our novelty is to jointly consider these variables and to pay attention to the robustness of the estimated elasticities and to their economic interpretation.

Our main conclusion is that augmenting the demand functions with non-economic geographically and timely detailed data does not substantially change the economic interpretation of demand models. Moreover, the penalized fit improves while an (albeit marginal) enhancement in the residuals autocorrelation and heteroscedasticity indicates more reliable estimates of price and income elasticities.

Additionally, previous literature mainly studies whole countries, regions, or small groups of countries, thus ignoring the issue of heterogeneity between (micro) destinations. Therefore, should different tourism segments have different elasticities, such studies would estimate averages that would mislead “micro” destinations’ policymaking. In our work we tackle the issue of heterogeneity by disentangling the lag/dependence structure and the GTs and CCIs structure that better detect specific origin/destination combinations.

Our findings suggest that there are indeed relevant differences between local destinations, likely stemming from the main type of tourism that is hosted (business vs. cultural vs. sea & sun) and from the intrinsic characteristics of areas. Similar considerations also apply to the differences found when alternative origin markets were

considered. This result is key from both the methodological and policy perspectives, as suggests the development and use of tailor-made models for each specific destination and the rejection of general models with common indicators and lag/dependency structures. The bottom line is that policy makers and researchers should adapt the baseline models to the specific combination of tourism, source market and destination under consideration, in order to capture the idiosyncratic characteristics of local tourism demand.

We now move to summarise the more specific results stemming from our case studies. First, we find positive and significant income elasticities, smaller in a “must see” cultural destination as Florence than in a Sea & Sun destination like Catania, perceived as a more luxury “extra” holiday. As expected, estimates of income elasticity are not significant for Milan, a business destination. Generally, the optimal lag structure for income is 2 months (UK in Catania and Florence, Germany in Florence), which is coherent with a choice pattern where the decision of travelling to such short-haul destinations is made with a short-lag. The optimal lag structure for Germany in Catania is 0 but the models with lag of 2 months however show very similar performances.

Second, relative prices, proxied by the ratio between CPI in origin and destination regions are generally found insignificant in Catania, the Sea & Sun destination where the cost of living is relatively low compared to European standards. The estimated price elasticities are instead significant (and negative, as expected) in more expensive destinations like Florence and Milan. Coefficients are more significant for the UK than for German demand: this confirms that – at least in “micro” destinations - tourists tend to be more aware of exchange rate changes rather than inflationary effects in the destination they plan to visit. No relevant difference is found between the models for arrivals and for overnight stays. Theoretically, the price level

should be more relevant in explaining the length of stay than the arrivals, but our case-studies were short-haul destinations, where the majority of trips are weekend breaks or short business trips.

Third, the inclusion of non-economic variables (GT and CCI) keep the sign and the significance of the economic coefficients (income and price) quite stable.

Nevertheless, moving from baseline to augmented specifications, we observe better overall fit and residuals statistics, suggesting a more reliable estimation of demand function's parameters. This feature is very important when policy makers' goal is to understand how sensitive different segments are to policy intervention. This is particularly true in the business travel destinations, where web search volume is often the only significant information, as climate indicators and price or income elasticities are rarely significant.

Fourth, in most of the specifications the sign of the GT variable is positive and significant at very short lags. Optimal lag is 0 (or 1, in the case of UK tourists in Catania), highlighting that tourists mainly search for information about the destination and its attractions in the proximity of the departure. The result does not exclude that tourists may use the Internet to decide among alternative destinations, but underlines the importance, for local stakeholders, of the web marketing activity in obtaining and sustaining competitive advantage.

Fifth, segments' heterogeneity also plays a key role that policy makers cannot underestimate. The best GT index to describe tourists demand dynamics changes with the origin/destination combination. Specifically, the category "Travel" is the most informative for Milan, a business destination, while the best indicator for Catania, a Sea & Sun destination, is "Beaches". The case of Florence, a cultural destination, is emblematic for heterogeneity as the fingerprints of the two origin markets are better

detected by the sub-categories “Travel” for Germans and “Hotel” and “Art” for British. It is hence straightforward to conclude that the high granularity of web-based information has to be exploited by econometric modelling and decision makers. On one hand, no simple recipe can be easily adapted to all markets but, on the other hand, the gain in terms of efficient use of information can be substantial.

Sixth, we find that CCI is significant (and with the expected positive sign) only in a few cases. In particular, it is never significant in Milan (as business tourism is less sensitive to climate) while in Florence and in Catania is significant only when demand is measured by arrivals. This is consistent with the assumption that climate conditions are linked to the decision-making process and to the trip motivations rather than to the length of stay. The optimal lag for CCI is seldom 12 (only for German tourists in Florence), the expected lag if tourists build their predictions on climate by looking at climatological statistics of the previous year. On the contrary, but consistently with Zhang & Kulendran (2017), the optimal lags for CCI to describe the demand dynamics often range from 1 to 3 months. The substantial coincidences with the lag of income and price variables led us to assume that tourists are less concerned with the comparison of temperature, precipitation and humidity across the seasons and more influenced by climate conditions, or by the climatological news when they book the journey.

This paper has many limitations, as the topic tackled in this paper is ever evolving: first, continuous improvements in the availability and accuracy of web-based geographically and timely detailed data call for further refinements in the model specifications that can be applied to study tourism demand in micro-destinations. **For example, patterns of micro-seasonality due to demand variations in different days of the week have not been addressed, although their relevance for local tourism policy is high: the use of daily GT and CCI data might shed further light on this issue, perhaps if linked**

to other data sources (e.g. accommodation prices provided by online search engines and occupancy rates provided by STR). Second, as the possible combinations between: the individual characteristics of the destinations, the several origin regions, and the different tourism products are many, our paper only analyses three destinations in one country: the application of our methodology to other case-studies might lead to the identification of robust general findings and shed further light on the role played by search data and climate indicators, which could be useful for both forecasting and policy-making. Finally, as a further extension of our approach, innovation in data collection and organization might imply that, in the near future, also official economic data (e.g. income and prices) will be published by National Statistical Offices with increased granularity, with possibility to include them in augmented models.

References

- Andergassen, R., Candela, G., Figini, P. (2013). An Economic Model for Tourism Destinations: Product Sophistication and Price Coordination, *Tourism Management*, 37:86-98
- Andergassen, R., Candela, G., Figini, P. (2017). The Management of Tourism Destinations: a Policy Game, *Tourism Economics*, 23(1):49-65
- Athanasopoulos, G., Hyndman, R.J., Song, S., Wu, D.C. (2010). The tourism forecasting competition. *International Journal of Forecasting*, 27, 822–844
- Bangwayo-Skeete, P.F., & Skeete, R.W. (2015). Can Google data improve the forecasting performance of tourist arrivals? Mixed-data Sampling Approach. *Tourism Management*, 46, 454-464
- Blake, A., & Cortes-Jiménez, I. (2007). The drivers of tourism demand in the UK. A report prepared by Christel De Haan Tourism and Travel Research Institute- University of Nottingham for the Department of Culture, Media and Sport. London: UK Parliament.
- Bornhorst, T., Ritchie, J. R., & Sheehan, L. (2010). Determinants of tourism success for DMOs & destinations: An empirical examination of stakeholders' perspectives. *Tourism Management*, 31(5), 572-589
- Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management*, 21(1), 97-116
- Choi, H., & Varian, H. (2012). Predicting the present with Google Trends. *Economic Record*, 88, 2-9
- Coshall, J.T. (2005). A selection strategy for modelling UK tourism flows by air to European destinations. *Tourism Economics*, 11 (2), 141–158
- Crouch, G.I. (1992). Effect of income and price on international tourism. *Annals of Tourism Research*, 19, 643-664
- Crouch, G.I. (1994). The study of international tourism demand: A review of findings. *Journal of Travel Research*, 33(1), 12-23
- Crouch, G.I. (1996). Demand elasticities in international marketing: A meta-analytical application to tourism. *Journal of Business Research*, 36:117–136
- Cuccia, T., Rizzo, I. (2011). Tourism seasonality in cultural destinations: empirical evidence from Sicily. *Tourism Management*, 32:589-595

- Dogru, T., Sirakaya-Turk, E., & Crouch, G. I. (2017). Remodelling international tourism demand: Old theory and new evidence. *Tourism management*, 60, 47-55
- Dwyer, L., & Forsyth, P. (2011). Methods of estimating destination price competitiveness: a case of horses for courses? *Current Issues in Tourism*, 14(8), 751-777
- Ettredge, M., Gerdes, J., & Karuga, G. (2005). Using web-based search data to predict macroeconomic statistics. *Communications of the ACM* 48(11), 87-92.
- Falk, M. (2013). Impact of long-term weather on domestic and foreign winter tourism demand. *International Journal of Tourism Research*, 15, 1-17
- Fuleky, P., Zhao, Q., & Bonham, C.S. (2014). estimating demand elasticities in non-stationary panels: the case of Hawaii tourism. *Annals of Tourism Research*, 44, 131–142
- Gil-Alana, L.A. (2005). Modelling international monthly arrivals using seasonal univariate long-memory processes. *Tourism Management*, 26, 867-78
- Goh, C. (2012). Exploring impact of climate on tourism demand. *Annals of Tourism Research*, 39(4), 1859-1883
- Goffi, G., & Cucculelli, M. (2018). Explaining tourism competitiveness in small and medium destinations: the Italian case. *Current Issues in Tourism*, 1-31
- Guizzardi, A. & Bernini, C. (2012). Measuring underreporting in accommodation statistics: Evidence from Italy', *Current Issues in Tourism*, 15 (6), 597–602.
- Gunter, U., & Onder, I. (2015). Forecasting international city tourism demand for Paris: accuracy of uni- and multivariate models employing monthly data. *Tourism Management*, 46, 123-135
- Gunter, U., & Onder, I. (2016). Forecasting city arrivals with Google Analytics. *Annals of Tourism Research*, 61, 199-212
- Haugland, S. A., Ness, H., Grønseth, B. O., & Aarstad, J. (2011). Development of tourism destinations. An integrated multilevel perspective. *Annals of Tourism Research*, 38(1), 268–290
- ISTAT (2018). Movimento turistico in Italia. Available at <https://www.istat.it/it/archivio/224376>
- Jeuring, J.H.G. (2017). Weather perceptions, holiday satisfaction and perceived attractiveness of domestic vacationing in The Netherlands. *Tourism Management*, vol. 61, 70-81

- Konovalova, A.A., & Vidishcheva, E.V. (2013). Elasticity of demand in tourism and hospitality. *European Journal of Economic Studies*, 4(2), 84-89
- Li, G., Law, R., Vu, H.Q., Rong, J., & Zhao, X.R. (2015). Identifying emerging hotel preferences using Emerging Pattern Mining technique. *Tourism Management*, 46, 311-321
- Li, H., Goh, C., Hung, K., & Chen, J.L. (2017a). Relative climate index and its effect on seasonal tourism demand. *Journal of Travel Research*, 1-15 (in press)
- Li, X., Pan, B., Law, R., and Huang, X. (2017b). Forecasting Tourism Demand with Composite Search Index. *Tourism Management*, 59, 57-66
- Lim, C. (1997). Review of International Tourism Demand Models. *Annals of Tourism Research*, 24(4), 835-849
- Mariani, M., Baggio, R., Fuchs, M., & Höepken, W. (2018). Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management*, 30(12), 3514-3554.
- Martins, L. F., Gan, Y., & Ferreira-Lopes, A. (2017). An empirical analysis of the influence of macroeconomic determinants on World tourism demand. *Tourism Management*, 61, 248-260
- Melotti, M. (2018). Florence: tourism, heritage and consumption. In M. Kozak, N. Kozak (eds.) *Tourist Behavior. An Experiential Perspective*, Springer
- Mietczkowski, Z. (1985). The tourism climatic index: A method of evaluating world climate for tourism. *The Canadian Geographer*, 29(3), 220-233
- Munoz, T. G. (2007). German demand for tourism in Spain. *Tourism Management*, 28(1), 12-22
- OECD (2011). *OECD Studies on tourism - Italy: Review of issues and policies*. Available at <https://www.oecd.org/italy/oecdstudiesontourism-italyreviewofissuesandpolicies.htm>
- Pan, B., Wu, D.C., and Song, H. (2012). Forecasting hotel room demand using search engine data. *Journal of Hospitality and Tourism Technology*, 3(3), 196-210
- Peng, B., Song, H., Crouch, G. I., & Witt, S. F. (2015). A meta-analysis of international tourism demand elasticities. *Journal of Travel Research*, 54(5), 611-633
- Presenza, A., Del Chiappa, G., & Sheehan, L. (2013). Residents' engagement and local tourism governance in maturing beach destinations. Evidence from an Italian case study. *Journal of Destination Marketing & Management*, 2(1), 22-30

- Rivera, R. (2016). A dynamic linear model to forecast hotel registrations in Puerto Rico using Google Trends data. *Tourism Management*, 57, 12-20.
- Schiff, A., and Becken, S. (2011). Demand elasticity estimates for New Zealand tourism. *Tourism Management*, 32(3), 564-575
- Sainaghi, R. (2006). From contents to processes: Versus a Dynamic Destination Management Model (DDMM). *Tourism Management*, 27(5), 1053-1063
- Sainaghi, R., & Baggio, R. (2014). Structural social capital and hotel performance: Is there a link? *International Journal of Hospitality Management*, 37(2), 99-110
- Sainaghi, R., & Baggio, R. (2017). Complexity traits and dynamics of tourism destinations. *Tourism Management*, 63, 368-382
- Sainaghi, R., & Mauri, A. (2018). The Milan World Expo 2015: hospitality operating performance and seasonality effects. *International Journal of Hospitality Management*, 72, 32-46
- Sainaghi, R., Mauri, A., & d'Angella, F. (2018). Decomposing seasonality in an urban destination: the case of Milan. *Current Issues in Tourism*, doi:10.1080/13683500.2018.143094
- Sainaghi, R., Mauri, A., Ivanov, S., & D'Angella, F. (2018). Mega events and seasonality: The case of the Milan World Expo 2015. *International Journal of Contemporary Hospitality Management*, 61-86
- Seetaram, N., Forsyth, P., and Dwyer, L. (2016). Measuring price elasticities of demand for outbound tourism using competitiveness indices. *Annals of Tourism Research*, 56, 65-79.
- Smeral, E. (2014). Forecasting the City Hotel Market. *Tourism Analysis*, 19(3), 339-349
- Smeral, E. (2017). Tourism forecasting performance considering the instability of demand elasticities. *Journal of Travel Research*, 56(7), 913-926
- Smith, P. (2016). Google's MIDAS touch: Predicting UK unemployment with internet search data. *Journal of Forecasting*, 35(3), 263-284.
- Song, H., & Li, G. (2008). Tourism demand modelling and forecasting - a review of recent research. *Tourism Management*, 29, 203-220
- Song, H., Witt, S.F., & Li, G. (2009). The advanced econometrics of tourism demand. New York, London. Routledge.
- Song, H., Li, G., Witt, S.F., & Fei, B. (2010). Tourism demand modelling and forecasting: how should demand be measured? *Tourism Economics*, 16(1):63-81

- Song, H., & Liu, H. (2017), Predicting tourist demand using big data, in X. Z. & F. D., Eds. *Analytics in Smart Tourism Design. Tourism on the Verge*. Springer Cham.
- UNWTO (2010). *World Tourism Barometer* (Madrid). Available at:
<http://www.unwto.org>
- Vu, J.C., & Turner, L.W. (2006). Regional data forecasting accuracy: the case of Thailand. *Journal of Travel Research*, 45, 186-193
- Witt, S., & Witt, C. 1995. Forecasting tourism demand: A review of empirical research. *International Journal of Forecasting*, 11, 447–475
- Yang, X., Pan, B., Evans, J.A., & B. Lv. (2015). Forecasting Chinese tourist volume with search engine data. *Tourism Management*, 46, 386-397
- Zhang, H.Q., & Kulendran, N. (2017). The impact of climate variables on seasonal variation in Hong Kong inbound tourism demand. *Journal of Travel Research*, 56(1), 94-107.

Table 1. Descriptive Statistics

Table 2. Best Baseline and Augmented models for Catania, optimal lags in parenthesis

Table 3. Best Baseline and Augmented models for Florence, optimal lags in parenthesis

Table 4. Best Baseline and Augmented models for Milan, optimal lags in parenthesis

	Germany			UK			Germany			UK			
	Mean	CV	D-H	Mean	CV	D-H	Mean	CV	D-H	Mean	CV	D-H	
Dependent variables							Google indices						
ArrCA	7.63	0.11	14.07**	6.68	0.10	9.35**	TravelCA	4.02	0.07	6.10*	3.93	0.09	5.04
ArrFL	9.31	0.08	49.86**	9.43	0.06	34.50**	HotelCA	2.99	0.19	12.41*	3.04	0.20	0.91
ArrMI	9.72	0.03	21.36**	9.65	0.03	4.59	ArtCA	2.75	0.25	15.75*	3.83	0.11	8.92
OverCA	8.75	0.08	10.22**	7.80	0.08	7.72*	BeachCA	3.60	0.13	1.07	3.40	0.14	12.31**
OverFL	10.64	0.09	50.78**	10.62	0.06	27.99**	TravelFL	4.02	0.07	1.43	3.83	0.11	1.78
OverMI	10.43	0.03	12.96**	10.41	0.02	2.60	HotelFL	3.84	0.09	0.51	3.67	0.13	3.66
Economic variables							ArtFL	3.37	0.16	31.86**	3.08	0.14	6.17*
Income	0.78	0.03	1.08	0.36	0.19	53.27**	BeachFL	2.96	0.23	1.54	2.75	0.28	0.65
PriceCA	-0.01	3.65	51.41**	-0.24	0.40	19.79**	TravelMI	4.14	0.05	2.35	3.99	0.07	2.71
PriceFL	0.01	3.56	12.91**	-0.23	0.38	14.47**	HotelMI	4.09	0.05	10.71**	3.97	0.07	0.56
PriceMI	-0.01	2.21	38.03**	-0.24	0.35	23.37**	ArtMI	3.12	0.17	2.48	3.24	0.17	3.62
Climate indices							BeachMI	1.91	0.36	9.35*	2.81	0.25	2.94
cci_ca_arr	0.45	0.18	20.4**	1.09	0.23	26.8**							
cci_ca_over	0.52	0.20	21.4**	1.16	0.23	26.7**							
cci_fi_arr	2.01	0.17	21.1**	1.45	0.18	19.8**							
cci_fi_over	1.95	0.18	20.8**	1.51	0.18	19.8**							
cci_mi_arr	1.99	-0.09	21.5**	1.14	-0.10	21.8**							
cci_mi_over	1.80	-0.11	21.5**	1.16	-0.10	21.9**							

Table 1. Descriptive Statistics

	Germany				UK			
	Arrivals		Overnights		Arrivals		Overnights	
	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>
C	0.065**	0.064**	0.067**	0.071**	0.156**	0.158**	0.165**	0.167**
Ar1	0.245**	0.259**	0.376**	0.369**	-	-	-	-
Ar12	-.0318**	-0.300**	-0.284**	-0.283**	-0.286**	-0.305**	-0.192*	-0.186*
Income	2,737**(0)	2.405**(0)	2.183*(0)	2.047*(0)	4.971**(2)	4.941**(2)	5.229**(2)	5.282**(0)
Price	-0.675 (0)	-0.613 (0)	-0.196 (0)	-0.286 (0)	-0.539 (2)	-0.253 (2)	0.142 (2)	0.586 (0)
Cci		0.176*(2)		0.073 (2)		0.097*(2)		0.067 (12)
Gtbeaches		0.091*(0)		0.096*(0)		0.191*(1)		0.217**(1)
AIC	6.704	-0.08	18.885	17.040	32.346	26.251	6.873	0.608
Normality	5.2424	4.5128	9.9079**	4.6531	59.849**	57.401**	0.3416	2.7306
Heterosced.	7.9904	8.0382	7.3199	8.6790	39.172**	42.799**	3.9772	2.3931
Autocorr.	0.9686	1.3101	1.9444*	1.7015	1.6971	1.5646	2.4393**	2.0951*

Table 2. Best Baseline and Augmented models for Catania, optimal lags in parenthesis

	Germany				UK			
	Arrival		Overnight		Arrival		Overnight	
	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>
C	0.021	0.036**	0.015	0.029	0.028**	0.036**	0.021*	0.035**
Ar1	-0.380**	-0.383**	-0.232**	-0.213**	0.424**	0.386**	0.565**	0.518**
Ar12	-0.473**	-0.422**	-0.422**	-0.399**	-0.187**	-0.213**	-0.124*	-0.144*
Income	1.653** (2)	1.739** (2)	2.316** (2)	2.329** (2)	1.637** (2)	2.124** (2)	1.205** (2)	1.382** (2)
Price	2.703* (2)	2.321* (2)	2.912 (2)	2.056 (2)	-0.329* (2)	-0.266* (2)	-0.253* (2)	-0.107 (2)
Cci		0.028* (12)		0.027 (1)		0.022** (3)		0.009 (3)
Gttravels		0.326** (0)		0.303* (0)				
Gtart						-0.0265 (2)		
Gt hotels								0.134* (0)
AIC	-145.249	-160.395	-56.654	-61.338	-222.597	-230.250	-230.066	-234.302
Normality	1.0781	0.1772	1.3889	1.6509	0.1855	0.3429	0.9244	1.7870
Heterosc.	10.1641*	14.6455*	5.9964	9.6859	2.0752	6.1561	2.5288	5.3736
Autocorr.	3.0748**	3.6114**	3.9083**	4.1050**	1.9695*	2.0230*	1.8739*	1.5796

Table 3. Best Baseline and Augmented models for Florence, optimal lags in parenthesis

	Germany				UK			
	Arrival		Overnight		Arrival		Overnight	
	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>	<i>Base</i>	<i>Augmented</i>
C	0.024*	0.046**	0.024*	0.045**	-0.024	-0.003	-0.023	-0.012
Ar1	0.181*	0.137	0.177*	0.161*	-	-	-	-
Ar12	-0.344**	-0.266**	-0.394**	-0.331**	-0.409**	-0.368**	-0.410**	-0.405**
Income	-0.766 (12)	-0.105 (12)	-0.722 (12)	-0.080 (12)	0.079 (2)	-0.040 (2)	0.402 (2)	-0.729 (12)
Price	0.565 (12)	2.087 (12)	0.843 (12)	2.249 (12)	-0.612**(2)	-0.537**(2)	-0.500**(2)	-0.592**(12)
Cci		0.021 (1)		0.024 (1)		0.058 (3)		0.042 (3)
Gttravels		0.421**(0)		0.405**(0)		0.204**(0)		0.239**(0)
AIC	-171.409	-184.792	-160.887	-171.841	-172.872	-177.429	-187.468	-192.548
Normality	2.5124	2.7129	4,9185	3.1517	8,5815*	5.8449	13,400**	16.1971**
Heterosc.	1.2028	4.3484	1,3120	5.7749	1,1858	1.4209	0.2227	5.1055
Autocorr.	2.3410*	2.3107*	1,8988*	2.1553*	0.8075	0.7852	1.2527	1.0960

Table 4. Best Baseline and Augmented models for Milan, optimal lags in parenthesis

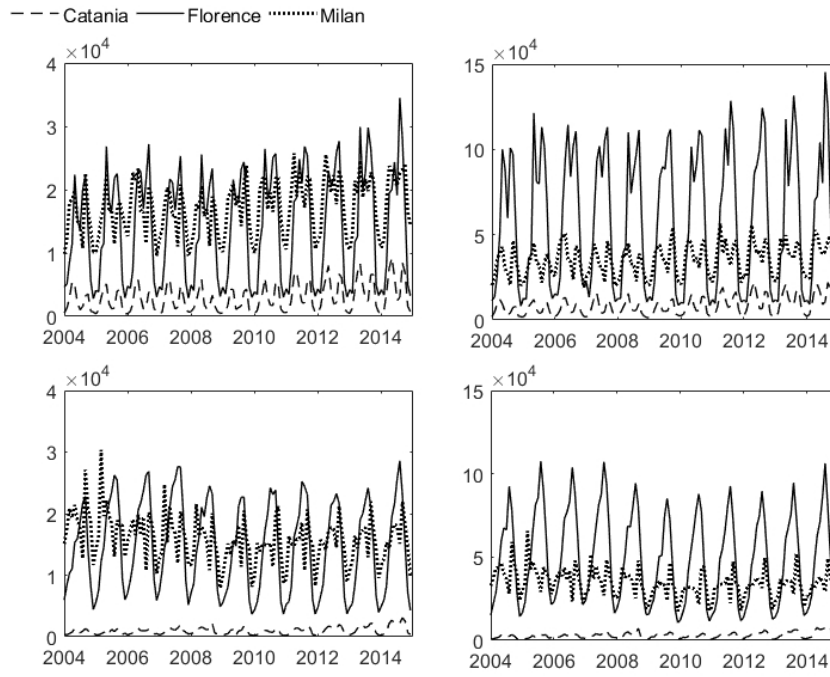


Figure 1. Arrivals (upper panels) Overnight stays (bottom panels) for Germany (lhs) and the UK (rhs)

221x151mm (96 x 96 DPI)

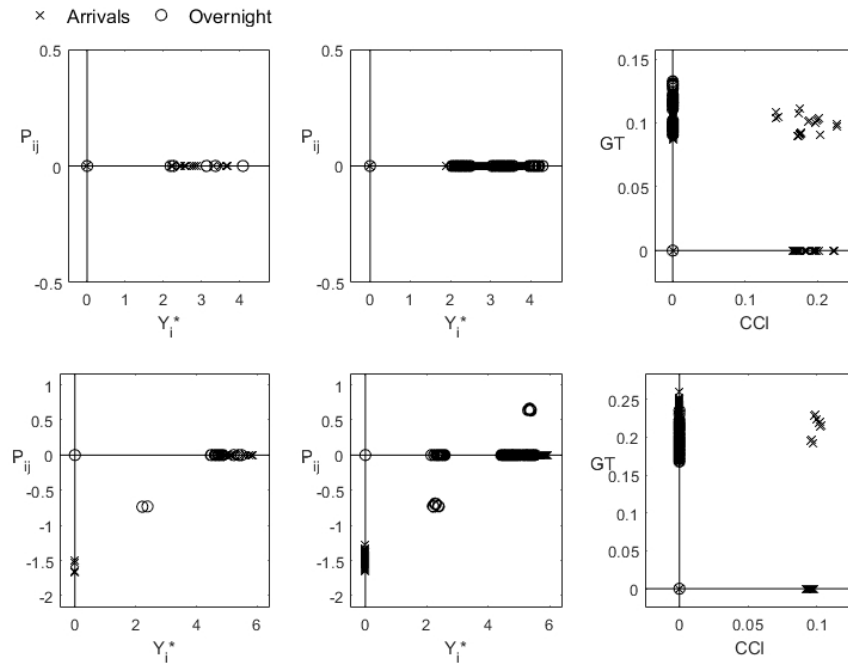


Figure 2. Estimated elasticities for Catania. Germany (upper panels), UK (lower panels). Baseline models (lhs), Augmented models (center- and rhs)

221x151mm (96 x 96 DPI)

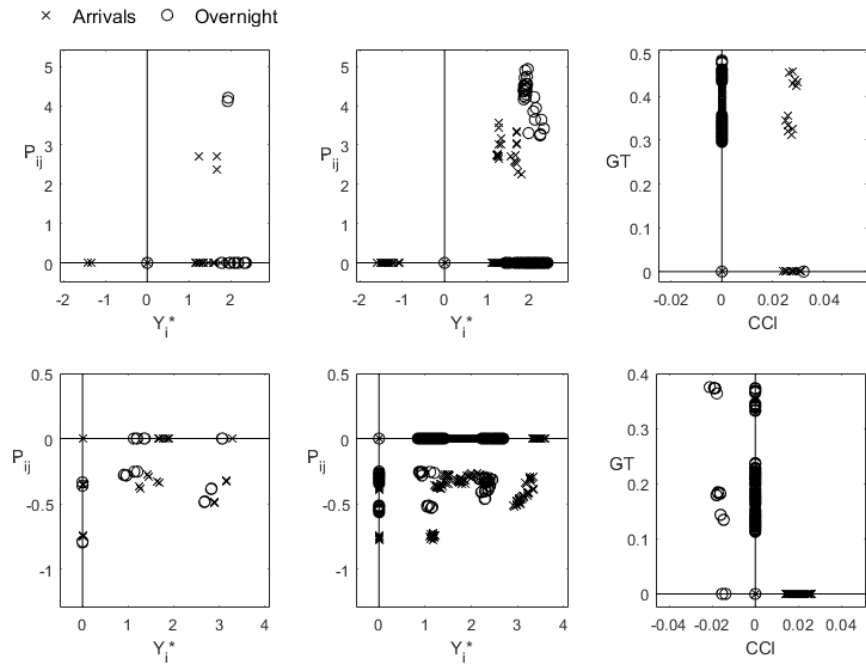


Figure 3. Estimated elasticities for Florence. Germany (upper panels), UK (lower panels). Baseline models (lhs), Augmented models (center- and rhs)

221x151mm (96 x 96 DPI)

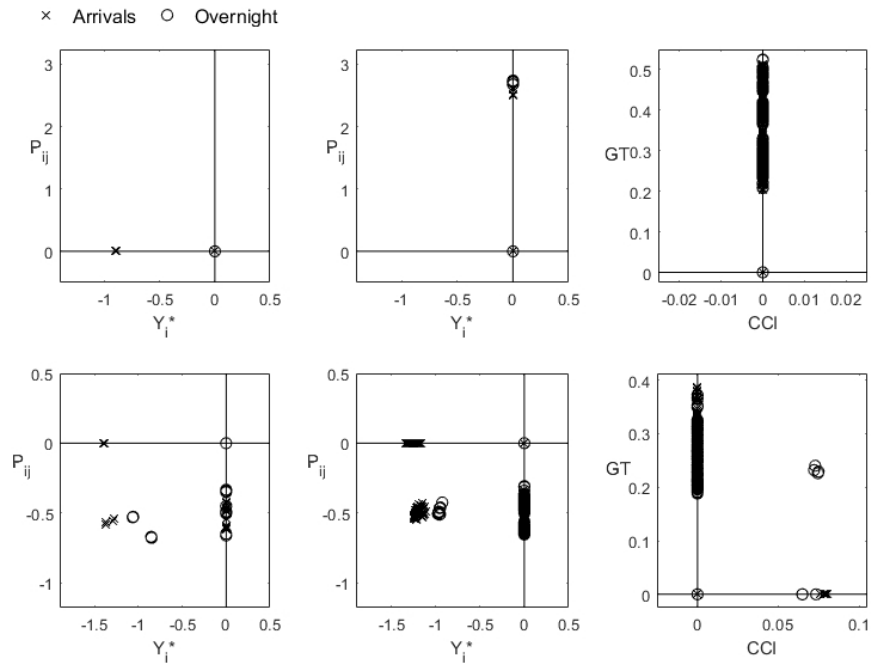


Figure 4. Estimated elasticities for Milan. Germany (upper panels), UK (lower panels). Baseline models (lhs), Augmented models (center- and rhs)

221x151mm (96 x 96 DPI)