Different tourists to different destinations.

Evidence from spatial interaction models

Emanuela Marrocu and Raffaele Paci

University of Cagliari, CRENoS

Abstract

As tourism is becoming one of the most important sources of economic growth at the local level, it is crucial to identify and assess the relevant determinants of tourism flows. This paper investigates this issue by carrying out an econometric analysis of tourism flows for the 107 Italian provinces based on the origin-destination (OD) spatial interaction models, which fully account for the spatial dependence generally featured by tourism flows. Besides geographical distance, the set of explanatory variables includes both pull and push locations' characteristics to assess their relative role in determining the distinctive traits of emissiveness and attractiveness for all the provinces. We thus consider income, density, accessibility, natural, cultural and recreational attractions.

The main results point out that there is a great deal of spatial dependence induced by neighbouring provinces at both origin and destination, which has been commonly overlooked by previous contributions relying on the gravity specification. Once we account for such a complex kind of dependence, we find sizeable impacts of neighbouring territories characteristics that enhance the effectiveness of internal provinces' tourism determinants. Most of the explanatory variables exhibit the expected effect, with distance and population density showing a negative impact on tourists' decisions when choosing a specific destination, while income, accessibility and attractions turning out to be crucial determinants of tourism flows.

Keywords: tourism flows, spatial origin-destination interaction models, neighbouring effects, product differentiation, amenities, Italy

JEL code: C21, D12, L83, Q26, R11

Acknowledgments: The research leading to these results has received funding from the Regione Autonoma Sardegna (LR7 2011, Project F71J11000980002). The authors would also like to thank Andrea Zara and Massimo Carboni for valuable assistance in preparing the database. We have benefited from valuable comments by participants to CRENoS conference.

Working Paper CRENoS 2012/10

1. Introduction

The tourism sector is becoming one of the most successful sources of economic growth at the local level. According to the World Travel and Tourism Council (2012a) the direct economic contribution of the tourism industry to the global economy has reached 2 trillion US dollars generating more than 100 million jobs. When the impacts induced in auxiliary sectors are also considered, the number of jobs increases to 260 million, yielding 9% of global GDP. The most relevant component of total tourism flows is the domestic one, which generated 70% of tourism GDP in 2011. In Europe, the holiday trips of EU residents in 2010 reached the remarkable number of more than one billion. As most Europeans spend their holiday trips in their own country, domestic tourism flows represent 77% of total trips (European Union, 2011). The tourism industry is also one of the most important economic activities in Italy; in 2010 total tourists' trips amounted to 96 million journeys and 366 million nights; also in this case the main component being the domestic holiday. The domestic and inbound tourism flows in Italy have generated 868,500 jobs directly and, considering also the other productive and service sectors related to hospitality activities, the tourism industry accounts for 2,231,500 jobs (9.7% of total employment) contributing to 8.6% of GDP (WTTC, 2012b).

It is thus important to identify and assess the relevant determinants of these rising tourism movements, which also have significant implications for decision-makers, economic operators and policy authorities alike, especially in destination locations.

The tourism literature has initially devoted more attention to the demand factors in explaining tourism flows between pairs of countries (see the survey by Lim, 1997); namely, income in the originating country, population, relative prices and geographical distance as measure of transport cost between the origin and destination countries. However, the demand approach overlooks the fact that the supply side - i.e. the tourism destinations - is extremely differentiated and, at the same time, consumers are characterised by heterogeneous preferences. Therefore, the diverse features of the leisure products play a key role in determining the flows of different tourists to different destinations (Smith, 1994; Papatheodorou, 2001).

The tourism market increasingly operates as a monopolistic competition market where destinations at different territorial levels (country, regions, cities) are horizontally differentiated and try to offer different varieties of leisure products which are effectively, or artificially, differentiated (Eilat & Einav, 2004) to heterogeneous consumers. This key aspect of supply differentiation has been extensively analysed by the literature on destination management; see, among many others, Dwyer, Forsyth & Rao (2000), Dwyer, Edwards, Mistilis, Roman & Scott (2009), Ritchie & Crouch

(2000) and Enright & Newton (2004). It has been remarked that tourists' evaluation is strongly connected to the complementary features of an overall integrated tourist product, like information and tourist services, cultural and natural resources, tourist safety (Cracolici & Nijkamp, 2008). At the same time, it is important to bear in mind that tourism destinations are vertically differentiated and provide different qualities of the leisure products. Quality differentiation is particularly influential in guiding tourists' choices of hotel accommodations (Skalpe & Sandvik, 2002).

In general, tourism flows can be considered as trade of services since they are equivalent to goods exporting activities for regions receiving the incoming tourists. Therefore, as in trade theory, both the demand and the supply side of the market must be considered; more specifically, in the case of tourism flows the demand factors are related to the origin, while the supply factors are linked to the destination.

The main purpose of this study is to examine the combined effects of demand and supply factors on the domestic tourism flows for the complete set of 107 Italian provinces using 2009 arrivals. We consider a highly disaggregated territorial level and this allows us to investigate in a more detailed way the characteristics of the destination area that are supposed to attract tourists. The explanatory variables, besides geographical distance, include both pull and push characteristics to assess locations' relative role in determining provinces' emissiveness and attractiveness traits. We thus consider income, density, accessibility, a set of cultural and natural endowments and other recreational attractions.

An important and original feature of our contribution is that we investigate tourism flows by carrying out an econometric analysis based on the recently proposed origin-destination (OD) spatial interaction models (Le Sage & Pace, 2008 and 2009). Such models permit to fully account for the spatial dependence, generally featured by tourism flows, which exhibits a quite complex pattern since tourists' movements are affected not only by geographical distance and by origin and destination specific features, but also by the characteristics of neighbouring locations at both origin and destination. Consequently, tourism flows may be simultaneously affected by two different kinds of spatial dependence, which so far have been neglected by most of the empirical tourism literature. One arises from the supply side and thus in the destination regions and the second is related to the demand side in the origin areas.

At the micro level this kind of dependence is likely to arise as the result of learning (at destination) and communication (at origin) processes as tourists share their travel experiences within family and friends networks. More specifically, it is likely that a tourist during her vacation in a certain destination also visits adjacent destinations, acquiring direct information on the neighbouring places which may became her tourist destination in future trips. In this respect destinations are seen as ‘experience suppliers’ (Ryan, 1997). Knowledge exchange processes related to tourism flows generating spatial productive spillovers in the destination regions are examined by Marrocu & Paci (2011), while Brau & Pinna (2012) provide evidence on the role played by tourists in triggering the imports of goods ‘experienced’ in visiting other countries.

As for the demand side, it is common that people talk about their vacations recommending their travel destinations to other potential tourists, such as friends or relatives in the places of origin. The information on the destinations spreads around the origin area, hence decreasing the uncertainty for potential visitors to that destination (Ryan, 1995). This generates a spatial spillover in the origin since it influences the propensity to travel to that destination among consumers in the areas contiguous to the origin. In the marketing literature repeated purchases or recommendations to other consumers are usually referred to as consumer loyalty and are seen as a key indicator of the marketing strategy success (Oppermann, 2000; Yoon & Uysal, 2005). Note also that consumers in the origin neighbouring regions are likely to share similar social and economic conditions and to have common preferences, this is particularly the case for homogenous groups, such as those of migrants. These kinds of similarities are likely to be taken into account by tourism operators in designing their offers to target customers in neighbouring provinces, and this in turn may induce spatial dependence. For Italy, this is particularly the case for the Northern areas of the country where immigrants from Southern regions are clustered according to a well-defined spatial pattern.

Moreover, the tourism industry, like many other production activities, is characterised by agglomeration externalities and this generates productive specialization patterns, especially for territories endowed with natural resources. Tourism activities tend to form territorial clusters which allow to share large sized infrastructures (airports, ports, conference centres, museums). This agglomeration process produces spatial dependence at the macro level and generates spatial spillovers across the destination provinces. Relevant examples in Italy of first tourist destinations which have generated the diffusion of tourism activities in contiguous areas are Rimini in the Adriatic coast and Costa Smeralda in Sardinia.

For all these reasons, the widely applied gravity model is underspecified as it relies just on a function of the OD distance to clear spatial correlation and on the bilateral features of the two areas involved in the flow of tourists. The existence of different kinds of spatial dependence is, on the contrary, consistent with a multilateral analysis perspective, which is expected to yield a more informative and comprehensive picture of the determinants shaping the economic links between the origin-destination pair of provinces since it accounts for the role played by their province neighbours.

Although this issue has been long acknowledged (Curry, 1972), especially for trade and migration flows, very limited evidence has been offered so far for tourism activities. The only exceptions are the very recent papers by Deng & Athanasopoulos (2011) and de la Mata & Llano-Verduras (2012). However, in both papers only the strength and significance of the spatial correlation is discussed, no measure is provided of how such correlation amplifies the impact of the internal tourism determinants of the geographical units analysed.

Therefore, considering the case of the Italian domestic tourism flows, the main original contribution of our paper is to assess empirically the relevance of both destination and origin spatial spillovers, along with their possible interaction, and to adequately estimate how the effects of the provinces' own internal determinants are enhanced by the positive influence of neighbouring areas.

The paper is organised as follows. In section 2 we briefly outline the recent literature on the determinants of tourism flows, focusing on the econometric contributions related to our work. Section 3 presents the main features of domestic tourism flows across the Italian provinces, while the selection of destination and origin determinants is discussed in section 4. The description of the methodology adopted to carry out the empirical analysis follows in section 5 as the selected econometric models and the preferred estimation strategy are strictly related to the previous contributions of the literature. They are also based on the features of the Italian domestic tourism flows and on the available data for the set of the explanatory variables. The econometric results are presented in section 6 and some concluding remarks are offered in section 7.

2. Related literature

The determinants of both domestic and international tourism flows have been extensively studied and the results are documented by a vast empirical literature (see the recent survey by Song, Dwyer, Li & Cao, 2012). As highlighted in Song & Li (2008), time series analyses received large attention as forecasting was the main focus of most studies, followed by the assessment of both supply and demand factors within multivariate regression frameworks. As might be expected, the evidence provided is quite differentiated given the wide range of methodologies adopted, the differences in time periods considered or in the coverage of geographical areas. A general and exhaustive survey of the previous empirical studies can be found in Crouch (1994), while specific contributions on the role of supply side differentiation are Murphy, Pritchard & Smith (2000) on destination marketing, Chen & Uysal (2002) on market positioning, Enright & Newton (2004) and Gomezelj & Mihalic (2008) on destination competitiveness.

In what follows we summarize the main results proposed in the most recent econometrics works, which serve as the basis to compare the findings of our analysis. In order to facilitate comparisons across the reviewed contributions, for each of them we report the main features of the analyses conducted, along with the most salient results in Table 1.

The international tourist flows are examined by Eilat & Einav (2004) who apply a multinomial logit model to an ample set of destination countries, which are considered as differentiated products suppliers. Having controlled for the relevance of trade flows between any countries' pair included in the sample, the presence of a common language and border, the climate and the perceived risk of the destination country, their findings point out that GDP elasticities are positive and statistically significant for both destination (0.81) and origin (1.29) countries. The latter result, with an elasticity above one, suggests that tourism is a luxury good. Moreover, the coefficient of geographical distance is negative (-0.98) signalling the effectiveness of transport costs in determining tourism flows. The important role played by transport infrastructure in influencing tourists' arrivals is confirmed by Khadaroo & Seetanah (2008), who examine bilateral tourism flows across 28 countries over the period 1990-2000. After controlling for consumers' persistence in the choice of the destination, they find that tourism flows are negatively affected by distance and prices, while their elasticity with respect to income in the origin country turns out to be positive (0.81), although lower than in other studies.

By adopting an international trade perspective, Zhang & Jensen (2007) test the hypothesis that supply side factors are the main determinants of international tourism arrivals; therefore, these are considered as a trade flow of services in the form of final consumers travelling to the destination countries to buy local products. In their regression analysis, they do not include any demand variable (like GDP or population size) and thus the potential for looking at tourism as a flow determined simultaneously by both origin and destination factors is somehow missed.

A different approach is followed by Garin-Muñoz (2009), who analyses the inflow of domestic and foreign tourists in a specific region, Galicia, during the period 1999-2006. Considering total nights spent, the estimated elasticities show that both domestic and foreign tourism flows are very sensitive to income in the origin markets and to prices. More specifically, in the case of foreign tourists both elasticities indicate a greater responsiveness due to the presence of a higher number of alternative destinations at international level.

Massidda & Etzo (2012) examine the determinants of the domestic tourism flows across the Italian regions over the period 2004-2007 within a Generalized method of moments panel estimation framework. The estimated elasticity for GDP in the origin region is equal to 1.42, indicating that domestic tourism has the connotation of a luxury good, with an impact that turns out to be much higher than that found in the previously cited studies. Tourism flows are also positively influenced by characteristics of the destination region like cultural expenditures, attractiveness, transport infrastructure and population density, confirming the crucial role played by the supply side factors. The negative impact of distance (-0.07) is also confirmed, but its magnitude is much smaller than in other studies. Finally, they found a negative impact of relative prices, but this result should be considered with caution since prices are measured by a consumer price index, which allows comparisons only across different periods in time for the same region and not across different regions.

As stated in the introduction, the issue of spatial dependence in tourism flows is directly tackled only in two very recent articles: Deng & Athanasopoulos (2011) and De la Mata & Llano-Verduras (2012).

Deng & Athanasopoulos (2011) propose a complex analysis of Australian domestic and international tourism flows, which is based on a dynamic spatial lag panel model applied to quarterly data for the period 1998-2008. The model accounts for both temporal and origin-destination spatial dependence that is also allowed to feature seasonal variation and asymmetry between capital-city and non-capital-city neighbours. Significant evidence of time-spatial correlation is found, along with positive effects of income and of time dummies controlling for two specific events, the Bali bombings and the Sidney Olympic Games. Note that no explicit measure of distance is included, even when the analysis is confined to only domestic flows.

De la Mata & Llano-Verduras (2012) analyse the domestic flows across the Spanish regions in two distinct years, 2001 and 2007, by using a Bayesian spatial autoregressive model. Although they find evidence of positive spatial autocorrelation, they do not model separately the spatial dependence which may affect in different ways the origin and the destination regions, as they impose from the beginning that the two spillover mechanisms are not separable. Instead of analysing the standard indicator represented by tourism arrivals, they compute a monetary measure of tourists' expenditure; surprisingly, GDP is introduced only for the destination regions, while the value added of the hotel industry and the beach length are included as explanatory features only for the origin. Their results confirm the negative influence of geographical distance (elasticity equal to -1.69).

Although both Deng & Athanasopoulos (2011) and De la Mata & Llano-Verduras (2012) present analyses based on spatial autoregressive specifications, they only report estimated coefficients for the explanatory variables and for the degree of spatial dependence without deriving to what extent spatial correlation enhances the equilibrium values of the tourism determinants, as recommended by LeSage & Pace (2009). In our analysis we will provide such kind of measures as they permit to assess the role of spatial spillovers and hence to distinguish the effects due to provinces' internal determinants from those generated by interactions among neighbours. This distinction is very relevant since policymakers and tourists operators should base their targets on a clear understanding of the relative strength of internal/external tourism flows determinants and strive to achieve the most favourable balance between the two in order to enhance territories' economic potential.

3. Domestic tourism flows in Italy

As we have seen in the introduction, the tourism industry in Italy represents one of the most relevant economic activities in terms of value added produced, employment and multiplier effects on several other manufacturing and services sectors. In 2010 Italy reached 96 million of total arrivals of tourists and 366 million of nights (Table 2). Over the period 2001-2010 all tourism indicators considered - domestic and international, arrivals and nights spent - have increased, showing a steeper trend for the foreign component and a decrease in the length of trips. Looking at the territorial breakdown, most arrivals are concentrated in the North of Italy (54%), whilst the South shows the longest average trip duration as the result of its specialisation in summer vacations. Domestic tourism represents the most relevant component in terms of both arrivals (54.9%) and nights (55.2%), although the domestic share has shown a slight decreasing trend in the last decade.

In this study we concentrate our attention on domestic tourism flows which, for a specific region, represent an important channel of inflow of external revenue. The Italian Statistical Office (ISTAT) publishes an annual report on the flows of domestic tourism disaggregated by province of destination and region of origin (ISTAT, 2009a).[1] In order to obtain the complete set of provincial bilateral flows, our first task is to estimate tourist flows by province of origin starting from the regional value. First, we used the annual survey on domestic vacation in Italy (ISTAT, 2009b) to compute the propensity to travel of the population in different age classes; secondly, we used the actual age distribution at provincial level to estimate the tourist flows by province of origin. This

[1] According to the Eurostat classification the 107 provinces in Italy correspond to the NUTS 3 territorial level and the 21 regions to the NUTS 2 level. It is important to remark that also the provinces in Italy have administrative authority especially for tourism activities.

way we obtain a square matrix for the 107 Italian provinces with 11449 Origin-Destination tourism flows in terms of domestic arrivals for 2009. Note that there are no cases of zero flows for the year 2009.

The geographical distribution of total domestic arrivals in the destination and origin provinces are depicted, respectively, in Map 1.A and Map 1.B, while the top ten provinces are listed in Table 3. Among the top ten provinces we find tourism destinations with well-defined specialisation and supply characteristics. These include large cities, like Milan, Rome, Turin, Naples and Florence, with world renowned cultural and historical attractions, but also destinations attractive for their natural endowments, like mountains (Bozen, Trento) or sea & sun (Rimini), or for religious sites (Perugia). If we look at the top provinces of outbound tourism flows it turns out, as expected, that among the top ten we find the Italian metropolitan areas characterised by the largest population (Rome, Milan, Naples, Turin) together with Northern provinces with high income levels. It is interesting to remark that some Southern provinces, characterized by a large population but a low income level (Palermo, Salerno, Catania), do not show high outward tourism flows.

This descriptive evidence suggests that some characteristics of the demand - and thus of the origin, like population and income - clearly affect the general propensity of outbound tourism flows. At the same time, other supply factors - like cultural, historical and natural amenities - are relevant in influencing tourism inflows.

However, as we have emphasized before, the distinctive characteristics of each bilateral OD flow need to be assessed in a framework allowing to account for the influence of neighbouring areas. For instance, in a 'sea & sun' destination like Rimini there is a high inflow of tourism arrivals (once normalised by population to get rid of the size effect) from mountain areas like Bozen, Trento and Aosta, while there are very low inflows from other nearby provinces characterised by similar coastal supply, as it is the case for Ancona or Pescara. Thus, demand and supply characteristics, as well as their spatial interaction, have to be considered as joint determinants of tourism flows. We provide a more general and rigorous evaluation of such spatial influences in discussing the econometric analysis results.

4. Destination and origin characteristics

As highlighted in reviewing the empirical literature, tourists' flows depend on a set of variables that accounts for various characteristics of both origin and destination areas in terms of economic, natural, cultural and territorial features. In this section, following the related studies presented in Table 1, we motivate our selection of the main explanatory variables included in the empirical models by distinguishing the factors related to the origin, the destination and those involving the origin-destination province pair. The complete list of all the variables considered as determinants of tourism flows, along with the data sources, is reported in Table 4.

Destination variables

GDP. The income level in the destination regions represents an indicator of the economic development in the receiving area and it is expected to influence the incoming tourism flows in two ways. First, a high income region provides better quality public services - like health care, public transport, law enforcement - which are important components of the product characteristics provided to tourists (Eilat & Einav, 2004; De la Mata & Llano-Verduras, 2012). Second, a high income area is more likely to attract business trips, which account for 25% of tourism spending in Italy. In both cases we expect a positive impact of GDP on tourism flows.

Density. One hypothesis often investigated in the tourism literature is the tourist behaviour toward the degree of congestion in the destination area (Massidda & Etzo, 2012; Saarinen, 2006). If tourists have preferences for less crowded areas, which may also signal a higher environmental quality, then we expect a negative sign. At the same time we may have the case of tourists who prefer more crowded locations associated with the presence of other attractions. This is consistent with the evidence reported in Santana-Jiménez & Hernández (2011) for the case of the Canary Islands tourism flows; they show that the effect of destination overcrowding can be different according to the tourists origin: Germans seem to like less crowded destinations while Britons are still attracted by high densities of people and do not perceive overcrowding in the islands. In our study, as a proxy for the degree of congestion of a tourism destination we consider the population density measured by the number of resident population per km^2.

Accessibility. An easy to reach destination certainly benefits from a factor which enhances tourism inflows. To measure this variable we have computed two alternative indicators. The first indicator is a five-group discrete variable that measures the potential accessibility of each province by road, train, air and time to the market (it takes values from 1=very low, up to 5=very high accessibility). The second one is the number of direct flights offered by low cost companies. For both indicators we expect a positive coefficient since higher accessibility should improve tourism inflows (Khadaroo & Seetanah, 2008; Massidda & Etzo, 2012). The number of direct flights is clearly a partial indicator of accessibility since the most common mode of transport for the Italian tourists is the car (62.9% of total trips in 2011, followed by the upward trending aircraft mode, 19.4%, ISTAT 2012). However, we have to bear in mind that in our models we also include the geographical distance in kilometres for each pair of provinces and thus, on the whole, the degree of accessibility by car is already taken into account.

Natural elements. Previous literature has shown how tourists are attracted by the natural environment of the destinations (Hughes, 2002; Brau, 2008), which we have measured by the number of protected natural areas located in the province; alternatively we have used the surface area of the natural parks in km^2.

Cultural attractions. An important pull factor of the destination is represented by cultural attractions like museums, churches and other monuments. Some of them are part of the cultural heritage built throughout several centuries of human development, while other cultural attractions - like museums designed by celebrated international architects - in most cases are the results of specific policies implemented to acquire a competitive advantage in supply side factors. For instance, this was the case for the Guggenheim museum in Bilbao (Plaza, 2000), which is one of the most widely known example of a modern cultural attraction not linked to the local historical heritage. Given the lack of a comprehensive database on cultural attractions like churches and monuments for the Italian provinces, in the empirical models we proxy such cultural attractions by the number of museums. Alternatively we have used the number of museum visitors, which may be seen as a more accurate indicator of the quality of the cultural attraction. As these attractions need some time for their effects on tourism flows to kick in, they are included in the empirical models with a two period lag, thus museum observations refer to the 2007 year.

Coast. Beach tourism is one of the most common tourism products in Italy; therefore, to assess the potential attractiveness of each destination for this widespread product, we have included a variable that measures the share of coastal municipalities over the province total.

Beach quality. The previous indicator measures the quantity of coastal supply but it is not adequate to account for its quality level. Therefore we have also included a variable that is supposed to capture the quality level of the coast, it is given by the number of beaches awarded the ‘bandiera blu’ (literally, ‘blue flag’) quality certificate by Legambiente (the top Italian environmentalist association).

Recreational attractions. To assess the role played by other types of leisure factors in attracting tourism flows - in addition to the ones related to natural and cultural elements - we have included the number of restaurants with at least 1 Michelin star at provincial level.[2] In general, the presence of restaurants with high reputation in a certain area signals the availability of high quality local products and of a widespread attention to the quality of life (Lambiri, Biagi & Royela, 2007) and this may represent a significant advantage to attract tourists with strong preferences for such factors (Sims, 2009).

Origin variables

GDP. Income in the origin region is one of the most important economic determinants in explaining tourist flows since it measures how the demand of consumers for travelling reacts to a change in their wealth. In general we expect a positive effect and it may turn out to be above one in case tourism is perceived as a luxury good.

Density. As we have discussed before, the population density is used to measure the degree of congestion of a specific province and we have thus included this factor among the origin variables to assess whether the overcrowding in the province of origin of the tourists' flows may affect their behaviour.

Origin-Destination variables

Geographical distance. In most analyses of domestic and international tourism flows the distance for each pair of destination and origin areas is included as a proxy of transport costs and thus it is expected to have a negative impact on tourism movements (Eilat & Einav, 2004; Khadaroo & Seetanah, 2008; De la Mata & Llano-Verduras, 2012). Note, however, that tourists preferences for specific destination's attributes (culture, climate, tranquility) may partly offset the negative influence of distance (Nicolau & Más, 2006). In this paper the geographical distance between each origin and each destination province is measured in kilometres.

Prices. Some studies include a measure of relative destination-to-origin prices (Eilat & Einav, 2004; Khadaroo & Seetanah, 2008; Massidda & Etzo, 2012). However, it is worth remarking that a price index for the tourism sector is not usually available - especially at the regional or provincial level - and, in general, it is not easy to define the price of a complex and

[2] Other kinds of activities - like leisure parks, disco or sport activities - may contribute to attract tourists' flows. Unfortunately, there are no official statistics available for these factors and it is not easy to collect homogeneous data at the provincial level.

differentiated product like tourism. Indeed, quality differentiation is very crucial in tourism supply and this makes the use of aggregate price index more problematic and less informative. Moreover, for the case of Italy it is not correct to use the general consumer price index (CPI) since ISTAT publishes price indexes which permit to measure price dynamics within a certain region but do not allow for cross sectional comparisons. Only recently, ISTAT (2008) published an experimental study on the price levels in the Italian regions for some specific products like food, clothes and furniture, which is supposed to allow for inter-regional evaluations. Although this price index may provide some cross-section indications of price variation, we are not convinced that it is able to capture the effective differences in tourism prices among destinations because it completely misses the accommodation costs, which is one of the most sizeable ones.[3] In any case, it is important to remark that transport costs, which are one of the largest components of total tourism costs, are already accounted for by the geographical distance. Moreover, given the high degree of differentiation of the tourism good, it is reasonable to assume that locations' attractions, income and transport costs receive prominent consideration in consumers' evaluations when choosing their holiday destination and that relative prices are assigned second order relevance.

5. Methodology

As anticipated in the previous sections the empirical analysis of the determinants of tourism flows is based on the application of spatial autoregressive models. Gravity models are also estimated as they constitute the usual starting point for analysing phenomena whose intensity and direction depends significantly, among other factors, on geographical distance. Moreover, this allows us to compare our findings with most of the existing empirical literature on tourism flows, which, so far, has largely relied on the gravity specification.

In this section we present the specific organization of the data adopted to handle the bilateral OD tourism flows and the most salient features of both gravity and spatial models (LeSage & Pace, 2008 and 2009).

[3] We have also considered an alternative source of data (the website of the municipality of Modena), which gives the price level of widely used products at the provincial level. In this case we have collected the average cost of a 'pizza and drink' meal that may reflect the average price of other products demanded by tourists.

5.1 Data organization

The bilateral provincial tourism flows from *n* origin provinces (with n=107) to each of the *n* destination ones are collected in the *n* by *n* square matrix **Y**. We consider Italy as a closed system of dimension $N=n^2$ and we adopt an origin-centric ordering, so that the columns of **Y** represent the origins and the rows the destinations. The diagonal entries of the matrix contain the intra-provincial flows, while the off-diagonal entries the inter-provincial ones. As the two kinds of flows may exhibit different characteristics and can be determined by different explanatory variables we deal with this issue by adopting the empirical specification that allows to account for the different variations in the two kinds of OD flows.

The dependent variable *Y* used in the econometric models is obtained by applying the *vec* matrix operator, which stacks the columns of the **Y** matrix into a column vector, $Y=vec(\mathbf{Y})$. Therefore *Y* has $N=n^2$=11449 observations: the first *n* observations pertain to the first origin province, the subsequent *n* observations to the second origin province, and so on.

The spatial distance between provinces is represented by the symmetric *n* by *n* matrix **G** whose entries are the geographical distance in kilometres between each origin and each destination province; in this case the main diagonal elements are set to zero. As was the case for the *Y* variable, the distance variable included in the estimated models, *G*, is obtained by stacking the columns of the **G** matrix ($G=vec(\mathbf{G})$).

The determinants of tourism flows are gathered in the matrix **X**, which has dimensions *n* by *k*, with *k* equal to 8, the number of the explanatory variables described in detail in the previous session. Given that the set of explanatory variables define both the attractiveness (when the province is considered as a destination) and the emissiveness (the province as an origin) features of the same province, the information contained in the **X** matrix is reorganized in three different matrices, X_d, X_o and X_{intra}. The X_d matrix is an *N* by *k* matrix obtained as $X_d=\iota_n \otimes X$ (where ι_n is a column vector of ones); the operator ⊗ is the Kronecker product which allows to repeat the entire initial **X** matrix *n* times (one for each of the *n* origin provinces) in order to represent the *k* provinces' characteristics at destination, so that they are seen as determinants of inbound tourism flows. Analogously, X_o ($X_o=X \otimes \iota_n$) is the *N* by *k* matrix representing the same characteristics, but this time seen from the origin standpoint, as they are meant to explain outbound flows. In this case the observations of the first origin province are repeated for the first *n* rows of the X_o matrix, the subsequent $n+1$ to $2n$ rows repeat for *n* times the features of the second origin province and so on. The X_{intra} matrix contains only information regarding the intra-provincial tourism flows.

5.2 Gravity model

The gravity model is specified as follows:

$$y = \alpha \iota_N + c\alpha_i + X_d\beta_d + X_o\beta_o + X_{intra}\beta_{intra} + \gamma\, g + u \qquad (1)$$

where the dependent variable y is the log-transformation of Y (Y=*vec*((**Y**)); $\alpha\iota_N$ is the intercept term. As described above, the matrices X_d, X_o and X_{intra} contain the provinces' characteristics when they are considered from a destination, origin or intra-province perspective, respectively. The vectors β_d, β_o and β_{intra} contain the corresponding effects. Note that the term $X_{intra}\beta_{intra}$ together with the intra-province flows intercept $c\alpha_i$ constitutes a separate model for the intra-provincial flows. In this way the parameters vectors β_d and β_o measure the exogenous variables effects on inter-provincial flows, which are the main focus of our study. Finally, the variable g represents the log of the geographical distance (G) between any two provinces in the origin-destination pairs. The error term, u, is assumed to be an i.i.d. process since, as stated in the introduction, the gravity model assumes independence among OD flows observations once the effect of distance is controlled for. However, such an assumption is unattainable when neighbouring provinces influence each other at origin, or at destination, or at both origin and destination, generating different and complex kinds of spatial dependence, which are properly dealt with within the spatial model framework, as described below.

5.3 Spatial autoregressive models

Following LeSage & Pace (2008, 2009), we tackle the tourism flows spatial dependence by considering the most general spatial autoregressive specification.[4] This amounts to augmenting the gravity model with three additional spatial autoregressive terms, based on connectivity matrices for destination, origin and origin-to-destination dependence; all three matrices are derived from the contiguity matrix. A given entry of the contiguity matrix assumes value 1 if the two provinces referring to that entry have a border in common, otherwise the value is set to zero. Before estimation the matrix is row-standardized and following conventional notation labelled as **W**.

The most general OD spatial interaction model is specified as a Spatial Autoregressive (SAR) model including three different spatial lags of the dependent variable:

[4] A comprehensive description of spatial models and related specifications, estimation and testing issues can be found in Le Sage & Pace (2009). Refer also to the contributions in the 2007 Journal of Econometrics special issue on spatially dependent data, edited by Baltagi, Kelejian & Prucha, and to the recent survey in Anselin (2010). For the case of individual data involving large sample sizes and complicated spatial weight matrices Bell &d Bockstael (2000) provide a useful application to household data of the Generalized Method of Moments proposed by Kelejian & Prucha (1999).

$$y = \alpha\iota_N + c\alpha_i + X_d\theta_d + X_o\theta_o + X_{intra}\theta_{intra} + \gamma\, g + \rho_d W_d y + \rho_o W_o y + \rho_w W_w y + \varepsilon \qquad (2)$$

where W_d ($W_d = I_n \otimes W$), W_o ($W_o = W \otimes I_n$) and W_W ($W_W = W_d W_o = W \otimes W$) are the three connectivity matrices described above; all the other terms are the same as in (1).

The spatially lagged term $W_d y$ is the weighted average of the tourism flows coming from neighbours of a destination province, so it is included in the model to capture destination kind of dependence. This arises when flows from an origin to a certain destination activate similar flows to neighbouring destinations; as claimed by Griffith & Jones (1980) the intensity of flows towards a destination are enhanced by the attractiveness degree of nearby destinations. The spatially lagged term $W_o y$, conversely, captures the dependence induced by the fact that the factors determining flows from an origin to a given destination may generate similar flows from neighbouring origins. The third spatially lagged term $W_w y$ is obtained as the interaction between the other two spatial lags as it is the average of flows from neighbours of the origin to neighbours of the destination and is named 'origin-to-destination' dependence by LeSage & Pace (2008).

As model (2) includes the three interaction terms, it allows to account for the fact that bilateral tourism flows do not depend only on the specific characteristics of the origin and the destination province and on their distance, but also on the features of their neighbouring provinces. Such kind of indirect interactions creates links among all the areas included in the system, which are neglected if one relies on the simple gravity specification.

Eight special cases of (2) can be obtained by imposing restrictions on the spatial autoregressive coefficients; when $\rho_d=\rho_o=\rho_w=0$, the gravity model results as a special case of (2); when $\rho_o=\rho_w=0$ or $\rho_d=\rho_w=0$ or $\rho_d=\rho_o=0$ only one kind of dependence is relevant, destination or origin or only their interaction. When $\rho_d=\rho_o$ and $\rho_w=0$ the model becomes a single weight matrix SAR specification, with the weight matrix given by $0.5(W_d+W_o)$ and an autoregressive coefficient equal to $2\rho_d=2\rho_o$, in this case it is assumed that the impacts of the X variables at origin and at destination are not separable and the model features a cumulative impact. This specification is the one adopted by De la Mata & Llano-Verduras (2012) to analyse the Spanish interregional monetary flows created by tourism activities. Note that all the restrictions on the autoregressive coefficients can be tested by means of likelihood ratio (LR) tests since the restricted models are all nested within the general model (2).

It is worth highlighting that the most important difference between the gravity model and the spatial autoregressive one is related to the interpretation of the explanatory variables effects. For the gravity model, as it is the case for any linear model, when a given explanatory variable (say x_r)

changes, its effect on the dependent variable is obtained by taking the partial derivative of y with respect to x_r, so that it is straightforwardly given by the estimated coefficient (say β_r).

In the case of spatial autoregressive models the existence of the spatially lagged terms - meant to capture the influences coming from neighbouring provinces - makes the computation of the effects more complex, but it has the advantage of making it possible to decompose the total effect of a given variable into a direct component, due to changes occurred in a province's own variable, and an indirect one, caused by changes in the same variable that take place in neighbouring provinces. As it will become clearer in presenting the empirical results, this distinction has relevant implications, both at the managerial and at the local government level.

More specifically, the complexity in deriving the effects arises because they have to be computed from the long-run relationship implied by (2), which is given by

$$y = (I_N - \rho_d W_d y - \rho_o W_o y - \rho_w W_w y)^{-1}(\alpha\iota_N + c\alpha_i + X_d\theta_d + X_o\theta_o + X_{intra}\theta_{intra} + \gamma\, g + \varepsilon) \quad (3)$$

where I_N is the identity matrix. The effects have to be interpreted as long-run equilibrium values, which result after all the interactions among the provinces have taken place across space and over time. In this framework the effect caused by a change in the x_r explanatory variable is no longer given by just the estimated coefficient associated with it (say θ_r), but it also depends on the ρs parameters, which measure the degree of the three kinds of spatial dependence, and on the spatial matrices, which capture the extent provinces are interconnected.

In summary, when a given variable changes, the SAR model allows us to know on average to what extent the effect on tourism flows is due to internal determinants (like natural and cultural attractions, tourism managerial abilities or local policymakers actions) or to the external factors activated by neighbouring territories, which are transmitted through the spatial interconnectivity structure. It has become conventional to report average measures of both the direct and indirect effects along with their dispersion measures.[5]

6. Econometric results

In this section we present the results of the empirical investigation on the determinants of domestic tourism flows in Italy.

In order to compare our findings with most of the existing empirical literature on tourism flows we firstly estimate models based on the gravity specification, which have been widely applied to investigate tourism flows, and then we turn to models based on the spatial specification which, as

[5] For a technical presentation on how such effects are derived refer to LeSage & Pace 2009 (chapter 2, pages 34-40, for a general treatment and chapter 8, pages 225-226, for origin-destination data models).

already emphasized, allow us to account for the complex pattern of neighbouring dependence featured by tourism flows, which so far have been largely overlooked in most contributions on tourism determinants.

In a preliminary investigation we estimate models including the complete set of explanatory variables at destination, origin and intra-province, in order to test whether it was statistically correct to exclude attractiveness indicators (accessibility, parks, museums, coasts and beach quality, recreational attractions) for origins and intra-provinces. As they turned out to be not significant at conventional levels, we include them only to characterize destination provinces, while GDP and density are included for all the three types of provinces. In this way we achieve a more parsimonious specification for our basic model.

In our preliminary analysis we also considered the role of prices by alternatively including the two indicators presented in section 4. However, they both show an unexpected positive sign, which can be reasonably attributed to the severe limitations that characterize the construction of these variables. Given the lack of reliable data on relative prices specific to the tourism sector and since we believe that a relevant component of total tourism costs - like transport costs - is already accounted for by the geographical distance, we prefer not to go further in dealing with this issue, so that prices are not included in the estimated models discussed below.

6.1 The gravity model

In Table 5 we present the results of the gravity model specifications. As we have remarked in the methodological section such results are likely to be unreliable as the gravity specification does not account for the spatial interactions among neighbouring provinces. By overlooking such interactions, the estimated effects have to be interpreted following a bilateral notion, so that they are entirely assigned to the internal factors of the pair of provinces, origin and destination, related to a given tourism flow. However, when spatial spillovers are indeed present this is likely to yield biased effects. For this reason the results reported below are interpreted at 'face value' only for the purpose of comparing them with previous contributions.

The estimated model (1) of Table 5 represents the basic specification, which, as stated above includes GDP and density at destination, origin and intra-province, while the attractiveness factors are included only as destination features. Most of the estimated coefficients in (1) are highly significant and they exhibit the expected sign; as for their magnitude, they compare favourably with the results provided by the recent empirical studies, summarized in Table 1. In particular, we find an elasticity of GDP which is just below 1 at origin or above 1 for intra-province flows, confirming that tourism can be considered a luxury good. The high elasticity for destination GDP signals that tourists' flows are enhanced by high levels of economic development and the availability of public services in the visited locations. Evidence of discouraging and adverse effects, on the other hand, is found for densely populated places, as flows are expected to decrease by 0.4% following a 1% increase in density. The origin-destination distance shows the expected negative effect on tourism flows with an estimated elasticity of -0.79, which is similar to the one reported in Eilat & Einav (2004), while it is much higher than the one provided by Massidda & Etzo (2012) for the Italian regions arrivals; on the other hand, the latter is exceptionally small when compared with the evidence reported in previous studies.

All the six different pull factors considered to assess the destinations attractiveness turn out to be quite effective in promoting tourism, although with different intensities. The lowest effect is found for cultural attractions (0.01%), proxied by the number of museum visitors recorded in previous years, whereas higher impacts are associated with the degree of accessibility (0.1%), the presence of parks (0.6%) and the provincial share of municipalities with coastal territories (1.36%). The impact of this indicator is reinforced by the presence of high quality beaches, which contributes with an additional effect of 2.8%. Finally, recreational attractions, proxied by Michelin restaurants, with an estimated effect of 2.2%, represent a very important determinant of tourism flows.

It is worth remarking that it is rarely the case that such a comprehensive set of pull factors is included in the analysis of domestic tourists' flows; we think that this is quite a novel and important contribution since it allows for a better understanding of the multiple aspects of a complex and highly diversified product like tourism.

In columns (2)-(4) we check whether the results of the basic model are robust with respect to the inclusion of alternative measures for accessibility, presence of parks and museums. The accessibility variable based on a comprehensive indicator of different transport modalities included in regression (1) turned out to be significant only at the 10% level, this is likely due to the fact that its effects might be to some extent picked up by the geographical distance. Therefore, in regression (2) we replace the transport multi-modality indicator with the number of low-cost flights. This is of course a less comprehensive measure of accessibility with respect to the previous one, but it has the advantage of providing additional and specific information on tourists' travel modalities over and above the one already included in the pure geographical distance proxy. As the low-cost flights variable exhibits a greater (0.5%) and highly significant impact on tourists' flows we prefer to keep such accessibility indicator in the subsequent estimated models.

The presence of park areas (regression 3), differently from the number of parks, induces a decrease of tourism flows assessed at about 0.02%; this may be interpreted as a sort of crowding-out effect, in provinces where there are large protected areas the unavailability of space for tourism activities more than offsets the positive effect of parks as a pulling factor. Finally, the number of museums turns out to be not significant (regression 4) and this highlights the finding that the attractive force is not the presence of museums per se but their quality, which is adequately represented by the number of visitors. Note that the estimated models (2)-(4) of Table 5 generally confirm the basic model results. Moreover, as far as multicollinearity is concerned, the computed variance inflation factors are well below the conventional threshold value of 6 signalling that collinearity does not affect the accuracy of our estimates.

6.2 Spatial autoregressive models

In the next step in our analysis of the Italian domestic tourists' flows we deal with the issue of spatial dependence by estimating four different specifications of the spatial autoregressive model. We adopt a specific-to-general approach by starting with the simplest SAR models, which account for one kind of spatial dependence at a time. In the first one we include the spatial lag of the dependent variable computed by relying on the destination weight matrix (W_d), while in the second model the dependent variable is obtained on the basis of the origin weights (W_o). The estimation results are reported in models 1 and 2 of Table 6; in order to save space we do not present the estimated coefficients, but the direct and the indirect effects along with their sum, the total effect.

As illustrated in the methodological section, the direct effect is due to changes in a given province's own explanatory variables, while the indirect one arises as the result of the interactions with neighbouring provinces. It is worth remarking that the estimation gains offered by the spatial specification over the gravity model are related to the significant existence of indirect effects, which amounts to properly accounting for cross-province spillovers. As illustrated below, all the spatial specifications we propose for the Italian domestic tourism flows yield highly significant and sizeable estimates for the indirect effects.

Focusing on the estimated total effects of the first model and comparing them with the basic gravity model results, it emerges that most of the explanatory variables' impacts are larger in absolute terms due to the existence of the spatial spillovers, estimated by the indirect effects. The only exception is represented by the OD distance; in the case of the gravity model its estimated coefficient is likely to be upward biased as it is not only capturing the effects genuinely associated with distance (such as transport costs) but also those arising from spatial interactions. It is also worth remarking that the estimated coefficient of the gravity model are quite similar to the direct effects obtained from the spatial specifications, this is especially the case for the destination SAR model; this is interpreted as an indication that the gravity model is too simple to be able to adequately capture the complex dependence structure featured by flows data and this results in unreliable estimates.

The relative intensity of both pull and push factors in driving tourist flows is strongly confirmed by the SAR models estimation, although the point estimates of the effects may differ substantially as they are a function of the kind of spatial dependence considered. Note that the strength of origin dependence turns out to be much higher than the destination one; the estimated spatial autoregressive coefficients $\rho_{o,d}$ are 0.69 and 0.34, respectively.

Having found significant evidence of origin and destination dependence when analysed individually, we also consider the empirical specification that encompasses both of them with the inclusion of the two spatial lagged terms for tourist flows described above. The results are reported in the third regression of Table 6, they offer further support to the relevance of both kinds of dependence and to the finding that interconnections at the origin are stronger than those at destination (the spatial autoregressive coefficient are estimated in 0.66 and 0.18). We tested the two-spatial-lag model against the model in which destination and origin weights are collapsed in one single matrix ($0.5W_d$+$0.5W_o$), as imposed in De La Mata & Llano-Verduras (2012), and the LR test (p-value 0.000) provided overwhelming evidence in favour of separable impacts due to destination dependence relations, on one hand, and to origin dependence, on the other; therefore, for the case of the Italian domestic tourists' flows we can confidently rule out the presence of a cumulative indistinct impact.

Focusing on the estimated effects, the destination-and-origin spatial model substantiates previous findings on the relative importance of the explanatory variables considered; it is worth noting that now spillover effects are predominant, accounting on average for around 80% of the total effect, due to the presence of both destination and origin sources of spatial interaction. Note that at this stage, however, spillover intensity is unlikely to be adequately measured, as we still have to account for the third possible type of spatial dependence, the origin-to-destination one. This is tackled by estimating the final model reported in regression 4 in Table 6; with respect to the previous SAR model, it comprises an additional lagged term of the dependent variable, which being computed on the basis of the interaction between origin and destination weight matrices (W_w) represents the average of flows from neighbours of the origin to neighbours at destination.

Featuring all three possible kinds of spatial dependence, the final model, as illustrated in section 4, is the most general one and on the basis of the LR test it is strongly preferred to all the other SAR specifications. Note that for this model the estimated strength of destination (ρ_d=0.62) and origin (ρ_o=0.86) dependence is much higher than was the case for the other spatial specifications, however it is partially offset by the origin-to-destination dependence which operates in the opposite direction (ρ_w=-0.65). In terms of the tourist flows this means that the effects of spatial spillovers obtained from the two previous spatial models were actually overestimated, as the complete set of spatial interactions which permeates the entire provincial system was underspecified. Moreover, as the most general specification accounts for the very complex dependence featured by flows data, total effects now have a multilateral interpretation. This happens because a change at origin (destination) of a push (pull) factor, through the interconnectivity structure, sets in motion a series of both push and pull events spreading across the entire system of provinces. Note that this kind of multilateral effects are completely missed within a gravity specification as the latter is meant to capture only the relative importance of the bilateral factors characterizing the link between the province of origin and the destination one.

According to the results reported in regression 4 spillovers effects are still quite sizeable, accounting for almost three quarters of the total effects. This is in line with the relevance of neighbouring features at both origin and destination provinces, which amplify the role played by bilateral characteristics. Focusing on total effects, income at the origin turns out to be quite effective in activating tourists' flows, the associated high elasticity of 2.2% confirms the luxury characteristic of the tourism product. The detrimental impact of crowded locations is confirmed as the density variable exhibits a negative effect. Once we account for all kinds of spatial interactions, which may run in opposite directions, the empirical results provide further support to the attractiveness traits of accessibility, natural resources and cultural amenities. More specifically, well preserved beaches are quite effective at attracting tourists' flows, as they may yield an increase of 3.7%, which reinforces the effect due to the presence of coastal areas (1%); renowned restaurants (recreational attractions) by signalling high life quality levels, may activate an additional 2.4% increase in arrivals, lower, but highly significant contributions in enhancing tourists' flows are due to accessibility (0.4%), parks (0.5%) and museums (0.01%).

When comparing the total effects obtained from model (4) of Table 6, it clearly emerges that the gravity model's results were very misleading. As a matter of fact, the latter were in magnitude similar to the spatial model (4) total effects, so that if one relies on the gravity specification she is led to assign the impacts on tourism flows *only* to the internal factors and to the assets of the pair of provinces sharing the same flow. This would be an incorrect conclusion as the greater part of the total effect is due to the existence of neighbouring spillovers and to their multiplier role captured by the significant spatial interaction terms. If such terms are not included the usual omitted variable estimation problem arises making the gravity estimates upward biased. This has important implications for both tourist operators and policy-makers alike. The distinction between the relative effect of internal and external determinants of tourism flows calls for effective coordination of managerial actions, destination marketing strategies and policies designed to promote tourism activities since they cannot be confined to the local level but need to be harmonized at the upper regional and national government levels.

Overall, our results offer sound empirical evidence on the most important driving forces of domestic tourists' flows in Italy. Besides geographical distance and income, a prominent role is played by the wide diversified set of locations' characteristics that provide a better understanding of some key aspects of the composite touristic good and may yield valuable indications on how the tourism sector may significantly contribute to local economic growth.

7. Conclusions

In this study we assess the most relevant determinants of domestic tourism flows for the Italian provinces by applying the recently proposed origin-destination spatial interaction models and simultaneously accounting for both demand and supply side factors. Although the issue investigated is rather relevant, as tourism is becoming one of the most successful sources of local growth, the existing economic literature, mostly focused on the demand factors, has devoted limited attention to the supply ones. These, however, are increasingly recognised to be key aspects of the 'tourism good' and they may have important implications for policies designed to promote long run sustainable growth by acquiring competitive advantages and making territories attractive to external consumers.

Tourism is a highly differentiated product, both horizontally and vertically, and destination places - aside from being greatly diversified among them - are featuring an ever more varied mix of characteristics to meet the preferences of highly heterogeneous consumers as tourists indeed are. For these reasons we analyse the combined effect of both demand and supply drivers of tourism flows in the whole set of 11449 bilateral flows for the 107 Italian provinces. The determinants considered thus include quite a comprehensive set of both pull and push location characteristics, namely income, density, accessibility, a set of cultural, natural and recreational endowments and geographical distance.

Differently from the traditionally applied gravity model, the application of the spatial autoregressive models has allowed us to tackle the issue related to the complex spatial dependence pattern exhibited by tourism flows. We find highly significant evidence on the existence of both origin and destination neighbouring provinces spillovers, which amplify the impact of internal determinants of tourists' flows. These findings offer further empirical support to the role played by a number of economic mechanisms and by consumers' behaviour, which work together to make provinces increasingly interdependent in the activities related to the hospitality industry. We recall that higher degrees of spatial interconnectivity in tourism flows are strictly related to the intensity of learning and communication processes, as tourists share their travel experiences within relatives and friends networks. This induces consumers' loyalty also in individuals at the origin who have not directly experienced the tourism product themselves and thus reinforces the links among origin locations. At the macroeconomic level the tourism sector is characterised by agglomeration externalities that favour productive specialization patterns and the formation of territorial clusters, which often extend far beyond the provincial boundaries, making profitable for destination locations to become more and more reliant on each other.

Focusing on the specific push and pull determinants, our results point out that the 'tourism good' has the connotations of a luxury good as income elasticity at origin turns out to be significantly higher than unity. Moreover, we find evidence that tourists' flows are enhanced by the existence of well-preserved beaches, parks, museums and renowned restaurants, while they are discouraged by overcrowding. Considering these elements as a whole, it seems that tourists are attracted by destination places showing a careful and caring attitude towards the environment and the cultural assets of the territory.

We think that these findings have relevant managerial and policy implications at both local and national levels. Tourism businesses could improve their perspective profitability by acting to make each destination a unique and versatile place to meet the very differentiated visitors' preferences. Strategic actions can be undertaken to promote the territory's appeal by increasing the number of cultural and recreational attractions, such as museums, events, exhibitions, shops and restaurants devoted to local culinary specialties, in traditional or new sophisticated fashion depending on the kind of targeted tourist customers. At the same time tourism business managers are expected to interact with local policymakers to ensure that the territory is adequately endowed with transport infrastructure and service facilities to enhance accessibility and holiday comfort.

Forward-looking policy-makers, being aware of the economic growth potential of the tourism industry, should envisage incentives schemes that, on one hand, encourage activities yielding economic value from the territory's assets, but on the other ensure a careful management of the natural environment, the artistic heritage and cultural resources in order to guarantee their preservation as an enduring source of growth. The relevant presence of spatial spillovers indicates that such policies cannot be effective if they are confined to the provincial level, but need to be coordinated at the upper regional and national levels. This kind of coordination, as a matter of fact, is already occurring among some national tourists operators and hotels chains, which are taking advantage of the origin-destination dependence and in so doing they are actually reinforcing it.

On the basis of the evidence found so far, it is our intention to extend the analysis in the future by addressing some limitations of the present study. In particular, we aim to investigate the determinants of tourists journey lengths, which are expected to provide a more accurate measure of the hospitality industry's contribution to the creation of national GDP when, as it is the case for Italy, data on tourists' expenditure are not supplied by national statistical offices on a regular basis, but can only be collected by means of specific surveys. Conditional on data availability, it would be also interesting to extend further the set of explanatory variables, in particular to account for the wide variety of cultural and leisure attractions, in order to gain a better understanding of the degree of attraction of the Italian provinces. Finally, insightful extensions regard the analysis of tourism flows over time and for specific subsamples. For example, it could be interesting to carry out the investigation for a particular destination to assess the specific determinants of its incoming tourists' flows in order to unveil possible comparative advantages with respect to the other provinces and, thus, provide a more accurate base of information to design businesses strategies and policies supporting economic specialization in the tourism sector.

References

Anselin, L. (2010) Thirty years of spatial econometrics, *Papers in Regional Science*, 89, 3-25.

Baltagi, B.H., Kelejian, H.H., & Prucha, I.R. editors (2007) Special issue on Analysis of spatially dependent data, *Journal of Econometrics*, 140, 1-332.

Bell P. K. & Bockstael E. N. (2000). Applying the Generalized-Moments Estimation Approach to Spatial Problems Involving Micro-Level Data, *Review of Economics and Statistics*, 82, 72-82.

Brau, R. (2008). Demand-driven sustainable tourism? A choice modelling analysis. *Tourism Economics*, 14, 691-708.

Brau, R., & Pinna, A.M. (2012). Moving of people for moving of goods? *The World Economy*, forthcoming.

Chen, J., & Uysal, M. (2002). Market Positioning Analysis: A Hybrid Approach, *Annals of Tourism Research*, 29, 987-1003.

Cracolici, M.F., & Nijkamp, P. (2008). The attractiveness and competitiveness of tourist destinations: A study of Southern Italian regions, *Tourism Management*, 30, 336-344.

Crouch, G.I. (1994). The study of international tourism demand: a review of findings, *Journal of Travel Research*, 33, 12-23.

Curry, L. (1972). A Spatial Analysis of Gravity Flows, *Regional Studies*, 6, 137-147.

De la Mata T., & Llano-Verduras, C. (2012). Spatial pattern and domestic tourism: An econometric analysis using inter-regional monetary flows by type of journey, *Papers in Regional Science*, 91, 437-470.

Deng, M., & Athanasopoulos, G. (2011). Modelling Australian domestic and international inbound travel: a spatial-temporal approach, *Tourism Management*, 32, 1075-1084.

Dwyer, L., Edwards, D., Mistilis, N., Roman, C., & Scott, N. (2009). Destination and enterprise management for a tourism future, *Tourism Management*, 30, 63-74.

Dwyer, L., Forsyth, P., & Rao, P. (2000). The price competitiveness of travel and tourism: a comparison of 19 destinations, *Tourism Management*, 21, 9-22.

Eilat, Y., & Einav, L. (2004). Determinants of international tourism: a three dimensional panel data analysis, *Applied Economics*, 36, 1315-1327.

Enright, M.J., & Newton, J. (2004). Tourism destination competitiveness: a quantitative approach, *Tourism Management*, 25, 777-788.

European Union (2011). *Statistics in focus*, 49.2011. Brussels: Eurostat.

Garin-Muñoz, T. (2009). Tourism in Galicia: domestic and foreign demand, *Tourism Economics*, 15, 753-769.

Gomezelj, D.O., & Mihalic, T. (2008). Destination competitiveness. Applying different models, the case of Sloveniam, *Tourism Management*, 29, 294-307.

Griffith, D.A., & Jones, K.G. (1980). Explorations into the Relationship Between Spatial Structure and Spatial Interaction, *Environment and Planning A*, 12, 187-201.

Hughes, G. (2002). Environmental indicators, *Annals of Tourism Research*, 29, 457-477.

ISTAT (2008). *Le differenze nel livello dei prezzi tra i capoluoghi delle regioni italiane per alcune tipologie di beni*. Roma: Istat.

ISTAT (2009a). *Movimento nelle strutture ricettive*. Roma: Istat.

ISTAT (2009b). *Indagine campionaria Viaggi e vacanze*. Roma: Istat.

ISTAT (2012). *Viaggi e vacanze in Italia e all'estero*, Statistiche Report. Roma: ISTAT.

Kelejian, H.H. & Prucha, I.R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model, *International Economic Review*,40, 509-533.

Khadaroo, J., & Seetanah, B. (2008). The role of transport infrastructure in international tourism development: a gravity model approach, *Tourism Management*, 29, 831-840.

Lambiri, D., Biagi, B., & Royela, V. (2007). Quality of life in the economic and urban literature, *Social Indicators Research*, 84, 1-25.

Le Sage, J.P., & Pace, R.K. (2008). Spatial econometric modelling of origin-destination flows. *Journal of Regional Science*, 48, 941-967.

Le Sage, J.P., & Pace, R.K. (2009). *Introduction to Spatial Econometrics*. Boca Raton: CRC.

Lim, C. (1997). Review of international tourism demand models, *Annals of Tourism Research*, 24, 835-849.

Marrocu, E., & Paci, R. (2011). They arrive with new information. Tourism flows and production efficiency in the European regions, *Tourism Management*, 32, 750-758.

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

Murphy, P., Pritchard, M., & Smith, B. (2000). The Destination Product and Its Impact on Traveller Perceptions, *Tourism Management*, 21,43-52.

Nicolau, J.L.& Más F.J. (2006). The influence of distance and prices on the choice of tourist destinations : The moderating role of motivations, *Tourism Management*, 27, 982-996.

Oppermann, M. (2000). Tourism destination loyalty. *Journal of Travel Research*, 39, 78–84.

Papatheodorou, A. (2001). Why People Travel to Different Places, *Annals of Tourism Research*, 28, 164-179.

Plaza, B. (2000). Guggenheim Museum's Effectiveness to Attract Tourism, *Annals of Tourism Research*, 27, 1055-1058.

Ritchie, J.R.B., & Crouch, G.I. (2000). The competitive destination, a sustainable perspective, *Tourism Management*, 21, 1-7.

Ryan, C. (1995). Learning About Tourists from Conversations: The Over-55s in Majorca, *Tourism Management*, 16, 207-16.

Ryan, C. (1997). *The tourist experience: A new introduction*. London: Cassell.

Saarinen, J. (2006). Traditions of sustainability in tourism studies, *Annals of Tourism Research*, 33, 1121-1140.

Santana-Jiménez, Y. & Hernández, J.M. (2011). Estimating the effect of overcrowding on tourist attraction: The case of Canary Islands, *Tourism Management*, 32, 415-425.

Sims, R. (2009). Food, Place and Authenticity: Local Food and the Sustainable Tourism experience, *Journal of Sustainable Tourism*, 17, 321-336.

Skalpe, O. & Sandvik, K. (2002). The Economics of Quality in the Hotel Business, *Tourism Economics*, 8, 361-376.

Smith, S. (1994). The Tourism Product, *Annals of Tourism Research*, 21, 582-595.

Song, H., & Li, G. (2008). Tourism demand modelling and forecasting. A review of recent research, *Tourism Management*, 29, 203-220.

Song, H., Dwyer, L., Li, G. & Cao, Z. (2012). Tourism economics research: A review and assessment, *Annals of Tourism Research*, 39, 1653-1682.

World Travel and Tourism Council (2012a) *World Economic Impact*. WTTC.

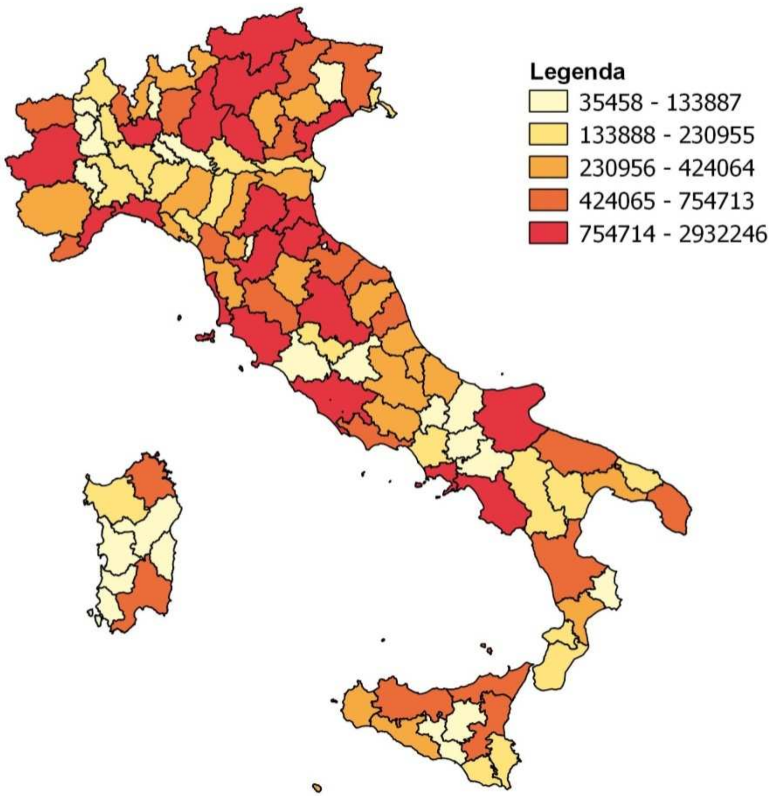
World Travel and Tourism Council (2012b) *Italy Economic Impact*. WTTC.

Yoon, Y., & Uysal, M. (2005). An Examination of the Effects of Motivation and Satisfaction on Destination Loyalty: A Structural Model, *Tourism Management*, 26, 45-56.

Zhang, J., & Jensen, C. (2007). Comparative advantage. Explaining tourism flows. *Annals of Tourism Research*, 34, 223-243.

Map 1. Domestic tourism flows (arrivals, 2009)

A. Province of destination



B. Province of origin

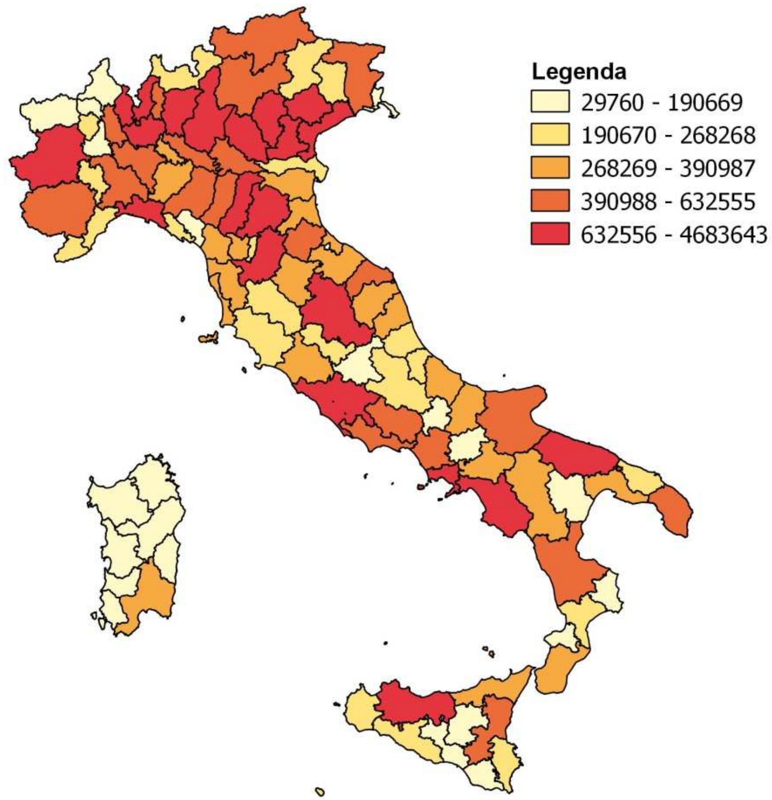


Table 1. Recent econometric studies on the determinants of tourism flows

Article	Period	Geo	Territory	Data	Method	Dependent variable	Tourists flow	Distance	GDP	Size	Relative prices	Other control variables
Eilat & Einav (2004, Tab 3. 3)	1985-1998	world	countries	panel	multinomial logit	arrivals	international	-0.98	O: 1.29 D: 0.81	O: no D: k: 0.62	-1.27	language, trade, border, climate, risk
Zhang & Jensen (2007, Tab 4.1)	1982-2001	world	101 countries	panel	two way fixed effect	arrivals	international	no	O: no D: 0.69	O: no D: p: 1.27	ns	hotel rooms, FDI, openess
Khadaroo & Seetanah (2008, Tab 2.1)	1990-2000	world	28 countries	panel	GMM	arrivals	international	-0.22	O: 0.81 D: no	O: no D: p: 0.30	-0.73	transport infr., hotel rooms, language, borders
Garin-Muñoz (2009, Tab 2)	1999-2006	Spain	Galicia	panel	GMM	nights	domestic	no	O: 0.86	no	-0.69	dummy for 2004 (holy year)
Deng & Athanasopoulos (2011, Tab 1)	1998-2008	Australia	83 statistical local areas	panel	dynamic OD spatial lag panel	nights	domestic	no	O: 19.4 D: ns	no	no	trend, capital-cities interacted with spatial terms, other dummies
de la Mata & Llano-Verduras (2012, Tab 8, M5_07)	2001, 2007	Spain	18 regions	cross section	bayesian spatial autoregressive	tourists expenditure	domestic	-1.69	O: no D: 0.78	no	no	islands, capital, beach, temperature
Massidda & Etzo (2012, Tab 6)	2004-2007	Italy	20 regions	panel	GMM	arrivals	domestic	-0.07	O: 1.42 D: no	O: d: 0.43 D: d: 0.71	-8.90	amenities, roads, crime, pollution

Note O: origin; D: destination; ns: not significant; no: not included. Size is measured in terms of either population (p), or density (d), or Km^2 (k)

Table 2. Tourism flows in Italy

A. Shares by macro areas (%)

	Arrivals		Nights	
	2001	2010	2001	2010
North	51.0	54.1	46.3	47.4
Centre	29.4	27.9	27.8	26.2
South	19.5	18.0	25.9	26.4
Italy (million)	*79.1*	*96.0*	*340.1*	*366.0*

B. Shares of domestic tourism over total (%)

	Arrivals		Nights	
	2001	2010	2001	2010
North	53.1	53.5	50.1	47.9
Centre	51.4	46.3	66.6	65.5
South	70.7	72.5	61.5	58.3
Italy	56.1	54.9	57.6	55.2

C. Annual average growth rate 2001-2010 (%)

	Arrivals		Nights	
	domestic	international	domestic	internation
North	3.3	3.1	0.6	1.7
Centre	0.4	3.0	0.0	0.5
South	1.6	0.6	0.4	2.1
Italy	2.1	2.7	0.3	1.5

Source: ISTAT

Table 3. Top ten provinces for destination and origin of domestic tourism flows
(million of arrivals)

Rank	Province of *destination*		Rank	Province of *origin*	
1	Milan	2.93	1	Milan	4.68
2	Rome	2.67	2	Rome	4.08
3	Rimini	2.41	3	Turin	2.50
4	Bozen	2.18	4	Naples	2.11
5	Venice	2.13	5	Brescia	1.48
6	Trento	1.98	6	Bergamo	1.30
7	Turin	1.65	7	Varese	1.04
8	Naples	1.52	8	Bari	1.02
9	Florence	1.25	9	Bologna	0.99
10	Perugia	1.24	10	Padova	0.91

Source: own calculation on ISTAT data

Table 4. Data sources and definitions

Variable	Definition	Primary source	Year
Tourism flows	arrivals to province of destination *i* from province of origin *j*	Istat	2009
Population	resident population (annual average)	Istat	2009
Density	population per Km^2	Istat	2009
GDP pc	GDP per capita	Istat	2008
Accessibility	potential accessibility by road, train, air and time to the market; five groups (from 1=very low, to 5=very high accessibility)	Espon	2006
alternative measure	low cost flights (number of direct destinations)	Companies web sites	2009
Parks	number of protected natural areas	www.parks.it	2009
alternative measure	size of protected natural areas (in km^2)	www.parks.it	2009
Museums	number of museums visitors	Istat	2007
alternative measure	number of museums	Istat	2007
Recreational attractions	number of restaurants with at least 1 Michelin star	www.michelin.it	2009
Coast	share of costal municipalities	Istat	
Beach quality	number of beaches with 'bandiera blu' quality certificate	www.legambiente.it	2009
Distance	distance in km between the provinces' centroids	Istat	
Prices	regional prices level	Istat	2008
alternative measure	average price of 'pizza & drink'	www.comune.modena.	2009

Table 5. Determinants of tourism flows. Estimated effects from gravity models
Dependent Variable: Tourism flows to Destination *i* from Origin *j*

	1	2	3	4
Destination				
GDP	1.048 ***	0.997 ***	1.057 ***	1.029 ***
Density	-0.390 ***	-0.366 ***	-0.431 ***	-0.381 ***
Accessibility: dummy	0.001 *			
Accessibility: flights		0.005 ***	0.005 ***	0.005 ***
Parks: num	0.006 ***	0.006 ***		0.005 ***
Parks: area			-0.016 ***	
Museums: visitors	0.009 ***	0.009 ***	0.006 ***	
Museums: num				-0.001
Recreational attractions	0.022 ***	0.023 ***	0.042 ***	0.024 ***
Coast	0.014 ***	0.013 ***	0.013 ***	0.013 ***
Beach quality	0.028 ***	0.038 ***	0.037 ***	0.040 ***
Origin				
GDP	0.917 ***	0.917 ***	0.917 **	0.916 ***
Density	0.066 ***	0.066 ***	0.066 ***	0.066 **
Intra province				
GDP	1.608 ***	1.608 ***	1.608 ***	1.608 ***
Density	-0.386 ***	-0.386 ***	-0.386 ***	-0.386 ***
Distance OD	-0.785 ***	-0.790 ***	-0.788 ***	-0.794 ***
Adj. R^2	0.722	0.724	0.721	0.723

Number of provinces: 107; total number of observations: 11449
Estimation method: OLS with White heteroskedasticity-consistent standard errors
Tourism flows, GDP, density, park area and number of museums visitors are log-transformed
All regressions include a constant and a dummy variable for intra-province flows
Level of significance: *** 1%, ** 5%, * 10%

Table 6. Determinants of tourism flows. Estimated effects from spatial autoregressive models

Dependent Variable: Tourism flows to Destination *i* from Origin *j*

	1			2			3			4		
Dependence	*destination*			*origin*			*destination* and *origin*			*destination, origin* and *origin-to-destination*		
Spatial autoregr. coef.	$\rho_d = 0.344$			$\rho_o = 0.692$			$\rho_d = 0.177$ $\rho_o = 0.665$			$\rho_d = 0.623$ $\rho_o = 0.864$ $\rho_W = -0.647$		
Effects	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total	Direct	Indirect	Total
Destination												
GDP	1.036	0.492	1.528	0.368	0.626	0.994	0.376	1.758	2.134	0.221	0.616	0.838
Density	-0.398	-0.189	-0.587	-0.128	-0.218	-0.347	-0.142	-0.663	-0.804	-0.077	-0.215	-0.292
Accessibility	0.007	0.003	0.011	0.002	0.003	0.005	0.003	0.013	0.016	0.001	0.003	0.004
Parks	0.005	0.002	0.007	0.002	0.004	0.006	0.001	0.007	0.008	0.001	0.004	0.005
Museums	0.011	0.005	0.016	0.004	0.007	0.010	0.005	0.021	0.026	0.002	0.006	0.007
Recreational attractions	0.023	0.011	0.034	0.010	0.017	0.026	0.009	0.043	0.052	0.006	0.018	0.024
Coast	0.014	0.007	0.021	0.004	0.007	0.012	0.005	0.023	0.028	0.003	0.007	0.010
Beach quality	0.032	0.015	0.047	0.016	0.027	0.042	0.012	0.054	0.066	0.010	0.027	0.037
Origin												
GDP	0.626	0.298	0.924	1.039	1.766	2.804	0.790	3.696	4.486	0.596	1.661	2.257
Density	0.046	0.022	0.068	-0.021	-0.036	-0.058	-0.029	-0.134	-0.162	-0.055	-0.154	-0.209
Intra province												
GDP	1.445	0.687	2.132	1.368	2.326	3.694	1.179	5.511	6.690	0.036	0.100	0.136
Density	-0.456	-0.217	-0.673	-0.324	-0.552	-0.876	-0.343	-1.602	-1.944	-0.163	-0.455	-0.618
Distance OD	-0.521	-0.248	-0.769	-0.223	-0.379	-0.602	-0.069	-0.322	-0.391	-0.190	-0.531	-0.721

Number of provinces: 107; total number of observations: 11449

Estimation method: ML for the first three models, Bayesian Markov Chain Monte Carlo for the last model

Tourism flows, GDP, density and number of museums visitors are log-transformed; accessibility is proxied by the number of low-cost flights

All regressions include a constant and a dummy variable for intra-province flows

All estimated effects are significant at the 1% level