sustainability

MDPI

Article

The Impact of High-Speed Rail Systems on Tourist Attractiveness in Italy: Regression Models and Numerical Results

Mariano Gallo [1,*] and Rosa Anna La Rocca [2]

1 Department of Engineering, University of Sannio, 82100 Benevento, Italy
2 Department of Civil, Building and Environmental Engineering, University of Naples 'Federico II', 80125 Napoli, Italy
* Correspondence: gallo@unisannio.it

Abstract: This paper evaluates the impact of high-speed rail systems on tourist attractiveness in Italy. The analysis is carried out with reference to provincial capitals, only some of which are served by high-speed railway lines. To achieve this objective, two multiple linear regression models were specified and calibrated, which relate arrivals and presences in accommodation facilities to several factors that could influence the tourist destination: cultural, historical, and monumental heritage, commercial activities, recreational activities, accessibility, etc. Both models showed that the availability of high-speed railway services is an important factor in the choice of tourist destination, being, moreover, the only accessibility variable found to be significant; furthermore, the elasticity of tourist demand to this factor was significant too.

Keywords: tourism; high-speed rail; accessibility; regression models

Citation: Gallo, M.; La Rocca, R.A. The Impact of High-Speed Rail Systems on Tourist Attractiveness in Italy: Regression Models and Numerical Results. *Sustainability* **2022**, *14*, 13818. https://doi.org/10.3390/su142113818

Academic Editors: Lóránt Dénes Dávid, Laszlo VASA and Setiawan Priatmoko

Received: 6 October 2022
Accepted: 23 October 2022
Published: 25 October 2022

1. Introduction

The accessibility of a site is one of the main factors influencing its ability to attract residential, commercial, and industrial settlements. In the literature, it is possible to find numerous works that confirm this assumption. Good accessibility makes a place easy to reach (passive accessibility) and allows those who live or work there to easily reach other places (active accessibility), making it more attractive to live in or establish a commercial or industrial activity.

From the tourism point of view, the accessibility of a site or a city can influence the choice of users when planning a trip, thus indirectly affecting the destination attractiveness; in fact, tourists tend to prefer an easily accessible destination over another that is more difficult to reach.

The accessibility of a place, in general, and even more so in tourism, is strongly influenced by the available public transport services. In this context, high-speed rail transport systems play an important role in the tourist development of a location, increasing its tourist accessibility as well as its general accessibility.

The purpose of this paper is to study the impact that high-speed rail (HSR) systems may have on tourism. To achieve this, two linear regression models were calibrated and specified to estimate tourist flows as a function of several accessibility variables, including the number of runs of high-speed rail services, as well as variables of cultural and tourist assets consistency. The models were calibrated with data from 111 provincial capitals in Italy, with reference to the year 2018, which is not affected by the impact of the COVID-19 pandemic. Although similar models could be calibrated for other Western countries, the Italian case study is significant because high-speed services are not widespread: of 111 provincial capitals, 49 are not served by high-speed rail at all, and 41 are served with no more than 10 rides per day.

In this study, we consider high-speed services not only those which exceed a maximum speed of 300 km/h (such as Trenitalia Frecciarossa services) but also those which reach a maximum speed of 250 km/h (Trenitalia Frecciargento services) and 200 km/h (Trenitalia Frecciabianca services), according to the UIC definition: "*High-speed rail combines many different elements which constitute a 'whole, integrated system': an infrastructure for new lines designed for speeds of 250 km/h and above; upgraded existing lines for speeds of up to 200 or even 220 km/h, including interconnecting lines between high-speed sections*" [1].

The limitations of the proposed models lie in their applicability only to the Italian case, but similar models can be specified and calibrated in other territories with the same approach proposed in the paper; from a temporal point of view, the models can be re-calibrated with reference to years different from the one under study, just as they can be applied after the construction of new high-speed railway lines to check whether the predictions remain valid. This work can contribute to the evaluation of investments in high-speed rail transport systems at the regional and national levels.

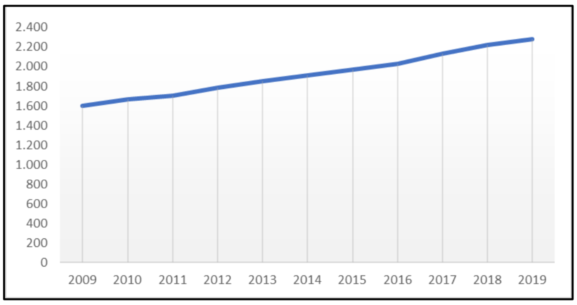
The goals of this study are basically twofold: (i) to verify whether and to what extent the presence of high-speed services has a real impact on tourist attractiveness; (ii) to carry out this verification through quantitative methods (mathematical models, in our case) that also provide a numerical estimate of the corresponding impact.

The paper is articulated as follows: Section 2 examines the background of the problem; Section 3 describes the data; multiple linear regression models are specified and calibrated in Section 4; an application to the city of Benevento is described in Section 5; conclusions and research perspectives are summarised in Section 6.

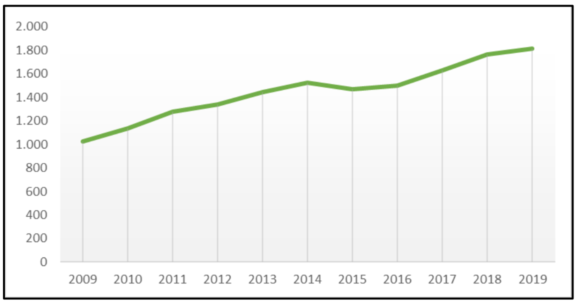
2. Background

2.1. Key Tourism Data

Over the past sixty years, tourism has steadily grown in both volume and importance, becoming one of the key pillars of the world economy. In 2018, for the ninth consecutive year, except for a period of crisis (2007–2009), international arrivals have grown (Figure 1).



Arrivals (billions)



Revenues (trillions)

Figure 1. Tourism arrivals (**left**) and revenues (**right**) in the world from 2009 to 2019 [2].

Data from the UN World Tourism Organisation [2] show that international tourist arrivals worldwide have reached 1.4 billion two years ahead of forecasts; it is due in part to the easiness of travelling, lower travel costs, the simplification of obtaining a visa, and some other factors that act as enablers for the expansion of tourism. Among destinations, Europe is the most popular, accounting for 51% of total arrivals in 2018 (710 million). Italy is the fifth destination in the top ten most visited countries in the world, with 62 million international arrivals (+7% from 2017 to 2018).

The data show that leisure activities and holidays are still the main purposes of the trip. The analysis of the purpose is useful for understanding the needs of tourism demand in terms of amenities and ancillary services that play a strategic role in the competition between tourist destinations. This is consistent with the traditional definitions of tourism: "*Tourism is a social, cultural and economic phenomenon which entails the movement of people to countries or places outside their usual environment for personal or business/professional purposes. These people are called visitors (which may be either tourists or excursionists; residents or non-residents) and tourism has to do with their activities, some of which involve tourism expenditure.*" [3,4].

Tourism is a crucial economic sector for Italy: in 2018, tourism had direct and indirect effects on GDP of 5% and 13.2%, respectively; furthermore, tourism directly generated 6% and indirectly 15% of total employment.

The attractiveness of 'destination Italy' is facilitated by the presence of a wide system of heritage attractions in terms of historical villages, archaeological sites, cities of art, cultural tradition, significant landscape, and seaside resorts. This system consists of 4026 museums, galleries, or collections, 570 monuments and monumental complexes, and 293 archaeological areas and parks. Moreover, 2371 municipalities host at least one museum, and there are 58 locations included in the UNESCO world heritage, positioning Italy at the top of the world ranking [5].

Analysing the movement of tourists by region, the data show that Veneto is the leader with over 19 million arrivals, followed by Lombardy with around 17 million; a second group includes Emilia Romagna, Tuscany, and Lazio, forming a central backbone throughout the country; Basilicata and Molise, on the other hand, fall into the lowest class of values (Figure 2).

2.2. Literature Review

In the state-of-the-art, we can identify three groups of works: (1) papers studying models for estimating tourism demand; (2) papers studying the impact of HSR services on aspects different from tourism; (3) papers studying the impact of HSR services on tourism.

Tourism demand estimation is a topic widely covered in the literature. A review on the subject can be found in [6].

There are several papers based on modelling time series and proposing forecast models based on them. Cho [7] compared three different approaches to forecasting tourist arrivals (exponential smoothing, univariate ARIMA, and Artificial Neural Networks, ANN) and found ANN to be the best method. Palmer et al. [8] proposed ANN-based models for forecasting tourism time series. Akin [9] proposed an approach for the selection of the best models for tourism demand estimation, starting from the comparison between three models used to determine time series; she defined a set of rules to identify the most suitable model according to the available data. Spatial interaction models for estimating tourism flows were proposed in [10]. Chan and Lim [11] analysed tourism seasonality in New Zealand using spectral analysis. An approach based on evolutionary fuzzy systems was proposed in [12]. Hassani et al. [13] proposed the use of Singular Spectrum Analysis (SSA) to predict tourist arrivals in the USA; the authors highlight that there are significant advantages of the proposed approach over more traditional ones such as ARIMA, exponential smoothing, and ANN. Li et al. [14] proposed a model based on principal component analysis and artificial neural networks for estimating tourism volumes based on time series. Other works on time series-based tourism flow forecasting can be found in [15,16], which proposed approaches based on structural time series, and [17], which also refers to stochastic nonstationary seasonality. Chu [18] proposed a fractionally integrated autoregressive moving average approach to forecasting tourism demand in Singapore. Andrawis et al. [19] applied time series to tourism in Egypt, while Nelson et al. [20] studied a case study in Hawaii.

There are many recent papers in the literature studying the impacts of HSR services from different perspectives. Cheng and Chen [21] studied the impacts on the capacity of traditional passenger and freight rail services.

Impacts on social exclusion/inclusion were studied by Dobruszkes et al. [22], who highlighted that the users of such services are predominantly "*...male, higher income, highly educated and belonging to higher social occupational groups*"; Ren et al. [23], who studied the impact of HSR services on social equity in China; and Cavallaro et al. [24], who studied the spatial and social equity aspects related to HSR lines in Northern Italy.

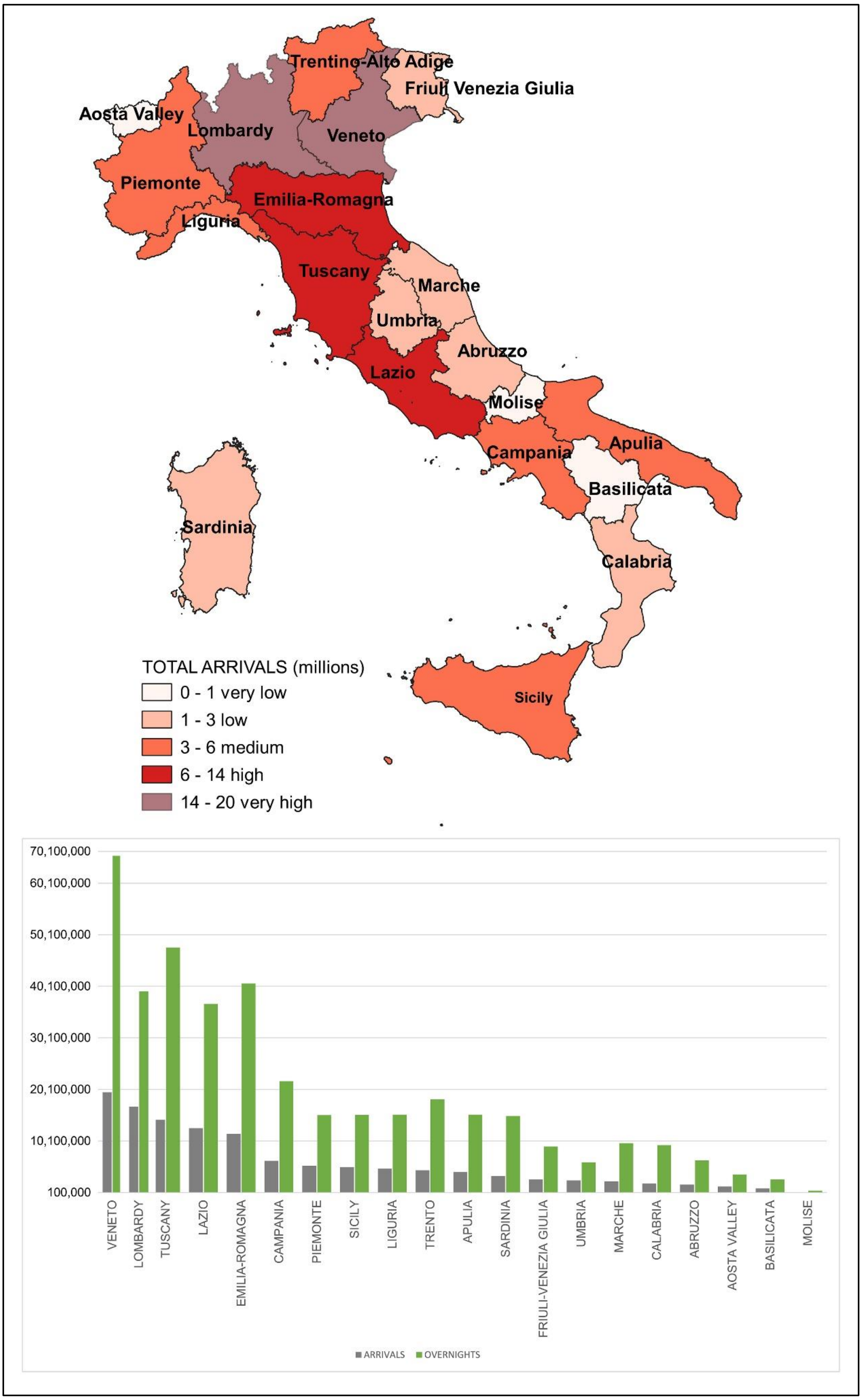


Figure 2. Movement of tourists by region.

Several papers have studied the impacts of pollution and/or greenhouse gas emissions, including the work of Fang [25], who studied the impact on air pollution in China, showing that it tends to decrease in regions where services are present, compared to unserved areas; Jia et al. [26], who studied the impact on CO_2 emissions in China, showing that there are significant reductions in greenhouse gas emissions; Strauss et al. [27], who studied the impact of HSR services on air transport demand and overall CO_2 emissions.

Several other papers have studied the impacts of HSR services on property values, including Huang and Du [28], who studied the effects on land prices in China, showing that the impact is significant, particularly in urban areas; Okamoto and Sato [29] examined the impacts of HSR services on land values, focusing on a region in Japan; Zhou and Zhang [30] studied the impacts on both property values and GDP.

The impact of HSR services on the industry was studied by Tian et al. [31], who examined the impact on service industry agglomeration in peripheral cities, showing that HSR facilitated economic growth in core cities at the expense of peripheral cities. The correlation between HSR services and the evolution of the high-tech industry in China was studied by Xiao and Lin [32], showing that the impact was significant, especially in cities. Chang et al. [33] studied the impact of the extension of the HSR network on industrial movement patterns in China's Greater Bay Area, showing that the expansion of the network led to a decentralisation of large industries. Zhang et al. [34] studied the impact of HSR on consumption in China.

Studies on the impact of HSR services on tourism are also numerous. In particular, the countries where these impacts have been most studied are China, Spain, and Italy. The main works concerning China were proposed by Wang et al. [35], who studied the effects of the HSR network on regional tourism development; Jin et al. [36], who studied the impact of HSR on winter tourism in a specific region; Wang et al. [37] studied the impact on urban tourism; another study was proposed by Yin et al. [38]. Zhang et al. [39] studied the impact of HSR on tourism mobility and the value of tourist firms. Zhou et al. [40] studied the effects of HSR on regional tourism economies in China. Campa et al. [41], on the other hand, studied the impact of HSR on tourism in both Spain and China.

The impacts of HSR services on tourism have been extensively studied in Spain. Pagliara et al. [42] studied the impact on tourism for the Madrid case study, while Albalate and Fageda [43] and Albalate et al. [44] studied the impact on tourism more generally. Two other works have been proposed by Guirao and Campa [45] and Guirao et al. [46].

The main studies referring to Italy are those by Pagliara et al. [47] and Pagliara and Mauriello [48], who studied the impact of HSR on tourism in Italy through statistical analysis.

Masson and Petiot [49] examined the impact on tourism attractiveness in a specific case study: the line between Perpignan (France) and Barcelona (Spain).

Other studies [50–52] investigated the potentialities of HSR for the tourism development of regions. Recently, also the Italian Minister of Culture Heritage indicated the use of HSR connections as a factor in revamping tourism after the pandemic event, especially in South Italy [53].

The research methodology adopted in this work involved the following main phases:

(1) Identification of the variables that can influence the choice of a tourist destination and data collection; in particular, three types of variables were identified: (a) variables related to tourism supply; (b) variables related to accessibility; (c) other variables that can influence the choice of destination (commercial activities, being a regional capital, etc.).
(2) Specification and calibration of multiple linear regression models capable of relating data on tourist attractiveness to the variables identified above. In this phase, the goodness of the models will be assessed through statistical tests, also to verify whether the assumption of linearity between dependent and independent variables is valid.
(3) Analysis of the results obtained and verification of the performance of the models by means of sensitivity analysis and an application to a case study.

These steps were preceded by a comprehensive analysis of the state-of-the-art.

In this paper, two multiple linear regression models are proposed to identify the main variables influencing tourist mobility in Italy; in addition to accessibility variables, including the presence of HSR services, data on the quantity of cultural heritage are considered in the model. As is shown in the following sections, of all parameters referring to accessibility, only the one related to HSR services was significant; other parameters, such as distances from airports or other municipalities, instead, were not statistically significant in explaining tourist flows.

To our best knowledge, the proposed approach was not proposed before in Italy, and similar studies are not available. Indeed, the studies available in the literature refer to the evolution of tourism following the implementation of new HSR services, but without providing models or quantitative methods capable of relating the variables of tourist attractiveness to the presence of rail links.

3. Data

Various data sources were used in this study. The main tourism data are taken from ISTAT (Italian Institute of Statistics) and quantify monthly arrivals and presences, classified by the origin and category of accommodation (hotel and non-hotel). These data were available at the regional, provincial, and municipal levels. Here, 'arrivals' correspond to the registration of customers in the accommodation facility, while 'presences' correspond to the total number of nights spent in a facility; therefore, in this study, the term 'presence' is equivalent to the term 'overnight stay'; in the following, we use the term 'presence' to be congruent with the ISTAT terminology. In the development of this work, the data on tourist movements refer to 2018, so that they are not affected by the COVID-19 pandemic event.

Overall, in Italy, there were 128.1 million arrivals and 428.8 million presences, with an average stay of 3.35 nights. The regional data on arrivals and presences are reported in Table 1, while Table 2 shows the same data with reference to the provinces of the regional capitals. It can be seen that Veneto, Lombardy, Tuscany, and Lazio are the regions with the most arrivals, while the provinces with the most arrivals are Rome, Venice, Milan, and Florence, with an obvious correlation with the attractiveness of the capital cities.

Table 1. Arrivals and presences in accommodation facilities in Italian regions [54].

Region	Arrivals	Region	Presences
Veneto	19,563,348	Veneto	69,229,094
Lombardy	16,757,628	Trentino-Alto Adige	51,416,000
Tuscany	14,188,009	Tuscany	47,618,085
Lazio	12,575,617	Emilia-Romagna	40,647,799
Trentino-Alto Adige	11,925,777	Lombardy	39,115,354
Emilia-Romagna	11,458,497	Lazio	36,684,847
Campania	6,234,863	Campania	21,689,412
Piemonte	5,276,117	Apulia	15,197,186
Sicily	4,998,055	Liguria	15,183,243
Liguria	4,718,832	Sicily	15,135,259
Apulia	4,065,979	Piemonte	15,100,768
Sardinia	3,280,894	Sardinia	14,940,111
Friuli-Venezia Giulia	2,610,097	Marche	9,656,538
Umbria	2,436,857	Calabria	9,277,810
Marche	2,256,564	Friuli-Venezia Giulia	9,022,550
Calabria	1,825,863	Abruzzo	6,335,072
Abruzzo	1,643,087	Umbria	5,937,298
Aosta Valley	1,254,191	Aosta Valley	3,606,289
Basilicata	892,087	Basilicata	2,603,622
Molise	138,570	Molise	448,600

Table 2. Arrivals and presences in accommodation facilities in main provinces [54].

Province	Arrivals	Province	Presences
Rome	11,131,197	Venice	36,628,413
Venice	9,677,150	Rome	32,245,018
Milan	7,718,958	Trento	18,156,000
Florence	5,245,117	Milan	15,717,859
Trento	4,415,851	Florence	15,281,325
Naples	4,149,784	Naples	14,199,255
Turin	2,505,985	Turin	7,248,575
Bologna	2,372,172	Perugia	5,099,833
Perugia	2,044,661	Bologna	4,729,192
Genoa	1,663,121	Genoa	4,055,435
Aosta	1,254,191	Aosta	3,606,289
Palermo	1,138,322	Palermo	3,286,743
Bari	1,096,477	Ancona	2,681,080
Ancona	754,777	Bari	2,475,938
Trieste	513,529	Catanzaro	1,524,800
Cagliari	460,221	Cagliari	1,463,800
L'Aquila	388,955	Trieste	1,188,103
Catanzaro	340,207	L'Aquila	919,851
Potenza	277,562	Potenza	743,220
Campobasso	101,579	Campobasso	363,210

On the other hand, the most attractive regions in terms of presence are Veneto, Trentino-Alto Adige, Tuscany, and Emilia-Romagna and the provinces Venice, Rome, Trento, and Milan. The difference between regional arrivals and presences, clearly linked to the average length of stay, is related to the type of holiday, often weekly, in Trentino-Alto Adige (mainly in winter periods) and Emilia-Romagna (mainly in summer periods).

Table 3 reports the data on arrivals and presences for the 111 Italian provincial capitals, on which the models have been specified and calibrated. The cities of Rome, Milan, and Venice have over 5 million arrivals and the same cities, with the addition of Florence, have over 10 million presences per year.

Table 3. Arrivals and presences in the provincial capitals [54].

City	Arrivals	Presences	City	Arrivals	Presences	City	Arrivals	Presences
Agrigento	172,984	267,359	Foggia	56,170	110,233	Pistoia	64,806	146,976
Alessandria	75,018	151,947	Forlì	108,429	224,361	Pordenone	58,238	135,653
Ancona	148,582	349,187	Frosinone	17,207	35,843	Potenza	37,526	66,897
Andria	19,045	33,947	Genoa	905,829	1,907,159	Prato	220,120	427,121
Aosta	96,542	190,356	Gorizia	29,987	69,111	Ragusa	156,569	514,334
Arezzo	229,458	434,418	Grosseto	249,998	1,120,975	Ravenna	621,961	2,744,504
Ascoli Pic.	35,946	78,574	Imperia	65,455	213,928	Reggio Cal.	77,631	186,424
Asti	46,461	95,349	Isernia	10,696	25,398	Reggio Emil.	199,571	367,647
Avellino	17,911	36,301	La Spezia	234,596	506,269	Rieti	28,383	68,260
Bari	446,394	838,627	L'Aquila	69,378	129,211	Rimini	1,856,268	7,460,300
Barletta	42,545	96,165	Latina	60,943	177,546	Rome	9,771,745	28,992,098
Belluno	60,526	169,303	Lecce	265,301	697,291	Rovigo	50,417	97,853
Benevento	36,252	80,144	Lecco	41,945	101,555	Salerno	252,455	621,362
Bergamo	350,418	636,535	Livorno	164,382	337,434	Sassari	70,487	144,404
Biella	44,465	110,423	Lodi	21,348	53,200	Savona	100,986	213,191
Bologna	1,543,053	3,059,546	Lucca	241,168	505,880	Siena	509,650	1,056,456
Bolzano	337,366	692,409	Macerata	40,962	221,122	Siracusa	253,732	749,719

Table 3. *Cont.*

City	Arrivals	Presences	City	Arrivals	Presences	City	Arrivals	Presences
Brescia	276,848	590,988	Mantua	124,472	214,375	Sondrio	14,434	25,645
Brindisi	77,227	159,399	Massa	192,668	834,724	Taranto	89,757	258,619
Cagliari	256,533	573,579	Matera	344,813	547,530	Teramo	21,373	64,449
Caltanissetta	27,164	104,567	Messina	36,342	83,604	Terni	95,998	209,827
Campobasso	13,203	23,237	Milan	5,695,214	12,058,835	Turin	1,290,390	3,800,003
Carbonia	11,024	31,997	Modena	270,411	571,425	Trani	48,023	94,761
Caserta	137,709	275,494	Monza	95,488	229,490	Trapani	72,211	173,769
Catania	474,025	975,888	Naples	1,376,589	3,684,905	Trento	360,388	1,016,951
Catanzaro	53,875	153,539	Novara	61,373	144,790	Treviso	159,924	332,341
Cesena	81,801	156,232	Nuoro	27,135	57,811	Trieste	414,003	929,492
Chieti	42,703	142,345	Oristano	68,568	147,137	Udine	210,598	389,112
Como	344,675	708,510	Padua	710,774	1,650,362	Urbino	77,951	487,446
Cosenza	58,823	105,306	Palermo	676,652	1,454,795	Varese	118,315	239,815
Cremona	70,569	136,761	Parma	386,160	735,127	Venice	5,255,499	12,118,298
Crotone	25,740	135,517	Pavia	57,577	110,439	Verbania	199,176	914,556
Cuneo	46,831	105,640	Perugia	425,875	959,070	Vercelli	25,626	82,678
Enna	36,529	62,901	Pesaro	228,445	776,171	Verona	1,198,279	2,495,943
Fermo	56,934	420,468	Pescara	143,025	256,164	Vibo Valen.	19,111	59,023
Ferrara	248,146	450,436	Piacenza	156,715	302,724	Vicenza	271,381	619,810
Florence	3,909,073	10,592,202	Pisa	782,288	1,882,097	Viterbo	88,661	201,896

We underline that the data used does not allow, at this territorial scale, to distinguish tourist trips from those for other reasons (work, business, study, etc.) and does not include stays in holiday homes or those trips that do not include a stay in an accommodation facility (one-day tourist visits, stays with relatives or friends, etc.); despite all these limitations, we believe that these data are the best available for the analyses we wish to conduct.

On the supply side, the accommodation establishments (see Tables 4 and 5) show the clear prevalence of Veneto and the Province of Venice, decidedly higher also than Lazio and the Province of Rome.

Table 4. Accommodation facilities in Italian regions [54].

Region	Accommodation Facilities
Veneto	72,363
Lazio	22,177
Emilia-Romagna	15,950
Tuscany	14,376
Trentino-Alto Adige	13,622
Lombardy	9845
Friuli-Venezia Giulia	7689
Apulia	7418
Campania	7185
Sicily	7155
Piemonte	7066
Marche	6935
Sardinia	5242
Liguria	5176
Umbria	4208
Calabria	3512
Abruzzo	3028
Basilicata	1409
Aosta Valley	1270
Molise	515

Table 5. Accommodation facilities in main provinces [54].

Province	Accommodation Facilities
Venice	41,906
Rome	19,126
Naples	3453
Perugia	3417
Trento	3330
Florence	3312
Bologna	2177
Milan	2068
Ancona	2049
Turin	2006
Bari	1723
Genoa	1299
Aosta	1270
Palermo	1192
L'Aquila	852
Trieste	832
Cagliari	735
Catanzaro	545
Potenza	544
Campobasso	374

Data on supply have not been used as possible explanatory variables in our models, since there is a direct relationship between supply and demand (supply increases where there is more demand) that could invalidate the modelling analysis aimed at identifying the other variables that can influence tourist flows.

Once the dependent variables had been identified and the corresponding data collected, the possible explanatory (or independent) variables were examined; these variables are the factors that could influence the choice of a touristic destination. Five categories of variables have been identified:

(a) Variables related to the supply of historical/cultural assets:
1. Number of state cultural sites [55];
2. Number of cultural heritage items [56];
3. Employees in libraries, archives, museums, and other cultural activities [57];
4. Consistency of historic urban fabric (elaboration on data) [58].

(b) Variables related to the supply of entertainment/amusement activities:
1. Employees in creative, artistic, and entertainment activities [57];
2. Employees in leisure and entertainment activities [57].

(c) Variables related to the supply of commercial activities:
1. Employees in retail trade (excluding motor vehicles and motorbikes) [57].

(d) Accessibility variables:
1. Number of direct runs on high-speed rail services (based on 2018 data);
2. Distance from Leonardo da Vinci airport in Rome (main Italian hub);
3. Distance from the nearest international airport;
4. Population-weighted road accessibility;
5. Total road travel time to all other possible destinations;
6. Total road travel distance to all other possible destinations.

(e) Importance variables:
1. Dummy variable (0/1) indicating the regional capital.

Not all variables refer to the same year. The most recent data on employees date back to the last census, which is carried out every ten years, but there are no better or more reliable statistical sources. On the other hand, the data on State places of culture and the stock of cultural assets, although referring to different years and before 2018, can be considered valid because the variation of these numbers over the years is negligible.

The following subsections describe the sources of the data and how they were obtained or derived.

3.1. Variables Related to the Supply of Historical/Cultural Assets

The number of cultural sites is a figure taken from [55] and refers to fortified architecture, archaeological areas, historical monuments, monuments of industrial archaeology, funerary monuments, archives and libraries, churches and places of worship, villas and palaces, archaeological parks, museums and galleries, parks and gardens. Only those under state jurisdiction and management are considered, and therefore, this variable does not include all possible cultural goods. This variable is indicated as scs_i, where i indicates the city.

The same source, but with reference to 2017 [56], provides the total number of cultural assets, understood as architectural assets, archaeological assets, parks, and gardens. This variable is indicated with tch_i.

The data on employees in libraries, archives, museums, and other cultural activities are taken from the ISTAT census [57]; clearly, the number of employees in this sector is assumed to be a proxy for the supply of the same type of activity to tourists. This variable is indicated with mus_i.

The size of the historical urban fabric was estimated from ISTAT data [58] by calculating the percentage of houses built before 1919. This variable is indicated with huc_i.

The values of these variables for the provincial capitals are shown in Table A1 in Appendix A.

3.2. Variables Related to the Supply of Entertainment/Amusement Activities

The data on employees in creative, artistic, and entertainment activities and employees in recreational and leisure activities are taken from the ISTAT census [57]. In addition, in this case, it is assumed that these data represent a proxy for the supply of this type of activity on the territory. The values for the provincial capitals are reported in Table A2 in Appendix A, and the variables are indicated, respectively, by ace_i and ree_i.

3.3. Variables Related to the Supply of Commercial Activities

The data on retail trade employees (excluding motor vehicles and motorbikes) are taken from the ISTAT census [57] and are assumed to be a proxy for the commercial offer in the territory. The values for the Provincial capitals are reported in Table A3 in Appendix A. This variable is indicated with ret_i.

3.4. Accessibility Variables

The tourist accessibility of a place, particularly a city, is determined by several factors depending on the infrastructures and transport services available. The data source or calculation methods for these variables are described below.

3.4.1. Number of Direct Runs on High-Speed Rail Services

This variable indicates the number of runs of Italian high-speed lines. The data refer to the number of runs of this type of service arriving/departing from the station of the municipality; for some municipalities, this value is zero, if not served by this type of service. This variable is indicated with hsr_i.

3.4.2. Distance from Rome's Leonardo da Vinci Airport (Italy's Main Hub)

The calculation of this variable, as well as all the following variables based on times or distances, required the construction of a graph of the national road network. This graph was implemented starting from the 'OpenStreetMap' database, correcting some connection errors and considering only the roads of the main network: all motorways; all primary roads with separated carriageways and their ramps; all main trunk roads (typically state roads and regional roads); some secondary roads necessary to ensure the full connection of the network.

Overall, this model represents 202,628 km of roads; Table 6 reports the extension of the network, while Figure 3 shows the overall graph. In addition to the length of the different road sections, which is necessary to calculate the distance between municipalities, it is also necessary to attribute a speed to each link, to calculate the corresponding travel time. In this work, we consider the use of the free-flow speeds sufficient, i.e., uncongested conditions, assuming the values reported in Table 7.

Table 6. Extension of the road network.

Type of Road	Total Length	Links
Motorways	14,592	23,463
Motorway ramps	2735	15,913
Primary roads	30,216	140,917
Primary road ramps	1403	16,717
Secondary roads	59,229	217,638
Secondary road ramps	432	7965
Tertiary roads	66,504	218,655
Tertiary road ramp	249	5852
Trunk roads	10,100	27,112
Trunk road ramps	2823	23,978
Other roads	14,345	98,456
Total	*202,628*	*796,674*

Table 7. Free-flow speed.

Type of Road	Free-Flow Speed [km/h]
Motorways	120
Motorway ramps	40
Primary roads	90
Primary road ramps	40
Secondary roads	70
Secondary road ramps	30
Tertiary roads	50
Tertiary road ramp	25
Trunk roads	70
Trunk road ramps	30
Other roads	40

With this model, the matrix of times and the matrix of distances between all the municipalities were generated; these matrices have a dimension of 8091 × 8091, being 8091 the Italian municipalities according to the 2011 ISTAT surveys. This matrix was simplified into a 111 × 8091 matrix, considering that the indicators were calculated only for the provincial capitals.

From this matrix, the variables in question were calculated as:

$$d^{AF}{}_i = 1/d_{i,Fiumicino} \qquad \forall i$$

where:

$d_{i,Fiumicino}$ is the distance between municipality i and Leonardo Da Vinci airport in Fiumicino (hundreds of km).

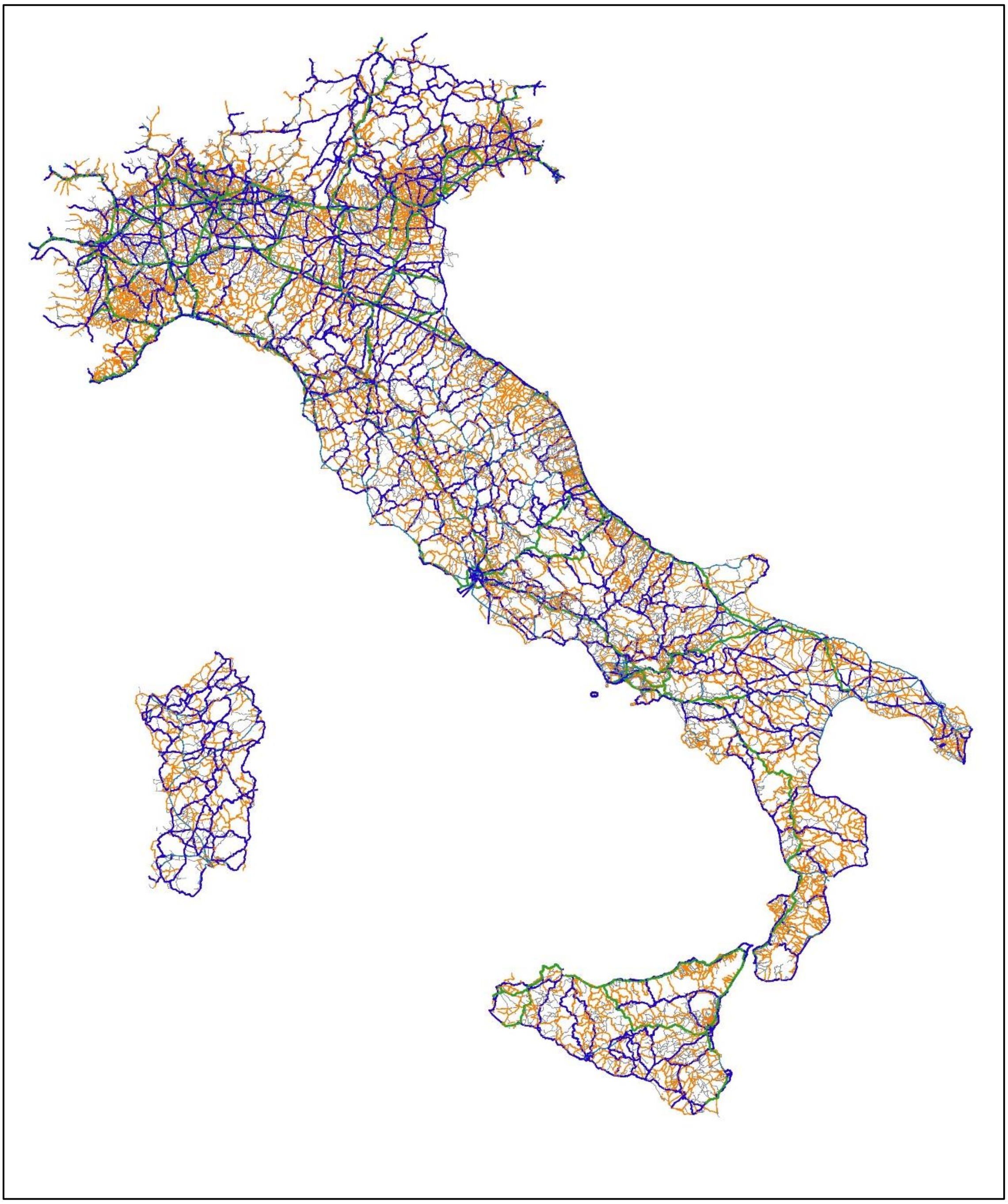


Figure 3. Road graph.

3.4.3. Distance from the Nearest International Airport

This variable was calculated as:

$$d^{AI}{}_{i} = \min\,(d_{i,AI}) \qquad \forall i$$

where:

$d_{i,AI}$ is the distance between municipality i and international airport AI (hundreds of km).

The international airports considered, those with the most traffic in each region, are listed in Table 8.

Table 8. International airports.

Region	Airport	Municipality
Piemonte	Turin–Caselle	Caselle Torinese
Aosta Valley	-	-
Lombardy	Milan–Malpensa	Ferno
Trentino-Alto Adige	Bolzano	Bolzano
Veneto	Venice–Tessera	Venice
Friuli-Venezia Giulia	Trieste–Ronchi dei Legionari	Ronchi dei Legionari
Liguria	Genoa–Sestri	Genoa
Emilia-Romagna	Bologna–Borgo Panigale	Bologna
Tuscany	Pisa–San Giusto	Pisa
Umbria	Perugia	Perugia
Marche	Ancona–Falconara	Falconara Marittima
Lazio	Rome–Fiumicino	Fiumicino
Abruzzo	Pescara	Pescara
Molise	-	-
Campania	Naples–Capodichino	Naples
Apulia	Bari–Palese	Bari
Basilicata	-	-
Calabria	Lamezia Terme	Lamezia Terme
Sicily	Catania–Fontanarossa	Catania
Sardinia	Cagliari–Elmas	Elmas

3.4.4. Population-Weighted Road Accessibility Function Variable

For the calculation of this variable, the ‘gravity-based measures’ model proposed by Hansen [59] was adopted. The general formulation of the model is as follows:

$$A_i = \sum_j W_j^{\beta} f(c_{i,j}, \alpha)$$

where:

A_i is the indicator measuring the accessibility of zone i;
W_j^{β} is a measure of the importance of zone j, based on activities, services, population, and so on;
β is a coefficient of the model;
$f(c_{i,j}, \alpha)$ is an impedance function, based on generalised cost, distance, etc., between zone i and zone j.

We have calculated the accessibility indicator as:

$$A_i = \sum_j inh_j \, t_{i,j}{}^{-1}$$

where:

inh_j is the number of inhabitants in the municipality j;
$t_{i,j}$ is the travel time in hours between municipality i and municipality j.

3.4.5. Total Travel Time by Road with All Other Possible Destinations

For each provincial capital, i, we calculated the total travel time (h $\times$ 10^{-3}) from all other Italian municipalities and calculated the variable as the reciprocal, with the following formula:

$$t^{tot}{}_i = 1/\sum_{\mathrm{j}} t_{i,j} \qquad \forall i$$

3.4.6. Total Road Distance to All Other Possible Destinations

For each provincial capital, i, we calculated the total travel distance (km $\times$ 10^{-5}) from all other Italian municipalities, based on the implemented graph, and calculated the variable as the reciprocal, with the following formula:

$$d^{tot}{}_i = 1/\sum_j d_{i,j} \qquad \forall i$$

where $d_{i,j}$ is the distance between capital city i and municipality j.

The values of all accessibility variables are reported in Table A4 in Appendix A.

3.5. Importance Variables

We consider a dummy variable indicating whether the city is a regional capital (1) or not (0). The values of these variables, indicated with cap_i, are reported in Table A5 in Appendix A.

4. Regression Models

The impact and significance of the explanatory variables on the tourism phenomenon are assessed with multiple linear regression models. These models relate the dependent variables (in our case, presences and arrivals) to the explanatory variables (independent) that may affect them.

Linear regression models take the following general form:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_k X_k + \ldots + \beta_m X_m$$

where:

Y is the expected value of the dependent variable;
β_0 is a coefficient of the model, which does not depend on the independent variables (intercept of the regression line);
β_k are the coefficients of the model, which together with $\beta_{0,}$ have to be calibrated;
X_k are the independent variables.

Any model must be specified and calibrated. The specification phase consists of defining which of the independent variables can be included in the model; the calibration phase consists of finding the coefficient values that can best reproduce the observed values of the independent variables for that specification.

The observed data of the independent variables are denoted by y_i and ordered in a vector $\boldsymbol{y}$; the vector $\boldsymbol{y}$ has as many elements as the number of municipalities on which we are going to calibrate the model (in our case, 111 municipalities). The values that the independent variables assume for each observation are also called 'predictors' and indicated with $x_{i,k}$, where i represents the provincial capital and k the independent variable; these values can be ordered in a matrix, $\boldsymbol{x}$, which has as many rows as the number of cities and as many columns as the number of independent variables plus one (coefficient β_0: the elements of the first column of the matrix are equal to 1). The coefficients β_k can be ordered in a vector $\boldsymbol{\beta}$ that has as many elements as the number of coefficients. Finally, we need to add the vector of statistical errors, $\boldsymbol{\varepsilon}$, which has as many elements ε_i as the number of cities. With these notations, it is possible to write:

$$y_i = \beta_0 + \sum_k \beta_k \, x_{i,k} + \varepsilon_i \qquad \forall i$$

or, in matrix form:

$$\boldsymbol{y} = \boldsymbol{x}\,\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

This formula represents, in short, the relationship between the observed data, $\boldsymbol{y}$, and the independent variables, $\boldsymbol{x}$. The calibration of the model consists in searching for the vector of coefficients, $\boldsymbol{\beta}$, that minimises the vector of statistical errors, $\boldsymbol{\varepsilon}$; in the theoretical case in which all statistical errors are equal to 0, the model would perfectly reproduce all the observed data.

If we denote by $\boldsymbol{x}^i$ the i-th row of the matrix $\boldsymbol{x}$, we can write:

$$y_i = \boldsymbol{x}^i\, \boldsymbol{\beta} + \varepsilon_i$$

whence:

$$\varepsilon_i = y_i - \boldsymbol{x}^i\, \boldsymbol{\beta}$$

The optimal values of the coefficients can be obtained using the generalised least squares method, which minimises the sum of squares of the statistical errors; the corresponding optimisation model can be written as follows:

$$\boldsymbol{\beta}^{\boldsymbol{opt}} = \text{Arg}_{\,\boldsymbol{\beta}} \min \textstyle\sum_i (y_i - \boldsymbol{x}^i\, \boldsymbol{\beta})^2$$

The ability of a model to reproduce observed data, and thus its goodness, is measured by several indicators; one of them is the coefficient of determination, R^2, which is calculated as:

$$R^2 = 1 - (\textstyle\sum_i (y_i - \boldsymbol{x}^i\, \boldsymbol{\beta})^2)/(\textstyle\sum_i (y_i - y\hat{})^2)$$

where $y\hat{}$ is the average of the y_i values; this indicator measures the ability of the y_i variables to explain the model, and the closer its value to 1 (statistical errors equal to 0 and perfect reproducibility of the observed phenomenon), the greater the goodness of the model.

The coefficient of determination always increases (or at least does not decrease) as the number of explanatory variables increases. To avoid this problem, it is possible to use the adjusted coefficient of determination, $R^2{}_{adj}$, which penalises the inclusion of variables that are not necessary to explain the phenomenon; this indicator is calculated as:

$$R^2{}_{adj} = 1 - ((n-1)/(n-p-1))\cdot(1 - R^2)$$

where n is the number of observations and p is the number of degrees of freedom (df) in the model. Clearly, as the number of explanatory variables, i.e., degrees of freedom, increases, the value of $R^2{}_{adj}$ decreases with respect to the value of R^2, the more so as there are few observations. In our case, with 111 observed data, we do not expect a great difference between the two values, which, in any case, will be calculated to verify the goodness of the model.

The coefficient of determination cannot, however, be the unique indicator to evaluate the goodness of a model. Indeed, it does not always decrease (it usually increases) with the number of variables k, even if some of them are not useful to explain the phenomenon. The other indicators that must be used to evaluate the model are the hypothesis tests that are able to measure whether the parameters adopted in the model are indeed significant to reproduce the phenomenon. In this study, we use the F-test, obtained from the analysis of variance, and the t-test, concerning the significance of each independent variable. We will assume that a model is acceptable if the significance F is close to 0 (at least < 0.05) and if the t-test of each coefficient β_k is higher [lower] than t_{95} [$-t_{95}$] for positive [negative] β_k, where t_{95} is the value of the t-student distribution corresponding to the degrees of freedom (df) of the model with 95% confidence. The degree of freedom of a model is equal to the number of independent variables x_k of the model. The values of t_{95} for the different degrees of freedom (1 to 10) are reported in Table 9.

Table 9. Values of t_{95} as a function of model degrees of freedom.

df	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
t_{95}	6.314	2.920	2.353	2.132	2.015	1.943	1.895	1.860	1.833	1.812

The specification and calibration procedure used in this study is based on a trial-and-error approach, based on the values of the Pearson correlation coefficients between the dependent variable and one of the independent variables. The correlation coefficient is calculated as the ratio of the covariance between two variables, σ_{xy}, and the product of the standard deviations, σ_x and σ_y:

$$\rho_{xy} = \sigma_{xy}/\sigma_x\sigma_y$$

This coefficient can assume values between -1 and 1; the higher the absolute value of the index, the more the two variables are correlated with each other, either positively or negatively, depending on the sign. The value of the correlation index indicates the possibility that the independent variable has a significant influence, within the model, on the dependent variable; therefore, in the trial-and-error procedure, variables with a higher absolute correlation index will be tested first, verifying if the sign is physically admissible. After a variable has been introduced, the model will be calibrated, and it will be checked whether the inserted variable is significant. If it is, the variable is kept in the model and another one is added; if it is not, another variable is tried. To be valid, a model must have all the independent variables significant, i.e., they must respect the minimum values of the indicator *t*-test, and a sign of the corresponding coefficient that has a physical meaning; among all the calibrated models that respect these conditions, those with the greatest coefficient of determination are preferable. This first phase leads to a model with all significant variables and with a coefficient of determination greater than all the other models tested; from this model, we try to introduce other variables and, then, to eliminate a variable and replace it with another, to test other possible combinations.

In Table 10, we report the correlation coefficients of each explanatory variable with the independent variables in decreasing order of value.

Table 10. Correlation coefficients.

Correlation Coefficients with the Independent Variable 'Arrivals'		**Correlation Coefficients with the Independent Variable 'Presences'**	
Variable	*Correlation coefficient*	*Variable*	*Correlation coefficient*
ace_i	0.896	ace_i	0.895
ree_i	0.862	ree_i	0.859
ret_i	0.853	scs_i	0.810
hsr_i	0.831	ret_i	0.807
tch_i	0.814	tch_i	0.802
scs_i	0.811	mus_i	0.795
mus_i	0.786	hsr_i	0.792
${d^{AF}}_i$	0.593	${d^{AF}}_i$	0.659
A_i	0.487	A_i	0.463
cap_i	0.459	cap_i	0.423
${d^{AI}}_i$	0.401	${d^{AI}}_i$	0.378
${t^{tot}}_i$	0.167	${t^{tot}}_i$	0.157
${d^{tot}}_i$	0.160	${d^{tot}}_i$	0.151
huc_i	0.042	huc_i	0.014

All specified and calibrated models are summarised in Table 11, for arrivals, and in Table 12, for presences, where, for each model, the considered variables, the R^2 and R^2adj indicators, the significance F, the model coefficients and, for each variable, the *t*-test value, whose limit value is also reported, and the validity or not of the model are indicated.

Table 11. Models to estimate arrivals.

Order	Variable\Model n.	1	2	3	4	5	6
1	ace_i	x	x	x	x	x	x
2	ree_i		x				
3	ret_i			x			
4	hsr_i				x	x	x
5	tch_i					x	x
6	scs_i						x
7	mus_i						
8	$d^{AF}{}_i$						
9	A_i						
10	cap_i						
11	$d^{AI}{}_i$						
12	$t^{tot}{}_i$						
13	$d^{tot}{}_i$						
14	huc_i						
	df	1	2	2	2	3	4
	R^2	0.802	0.805	0.809	0.843	0.882	0.882
	$R^2{}_{adj}$	0.800	0.801	0.806	0.841	0.879	0.878
	Significance F	3.95×10^{-40}	5.26×10^{-39}	1.41×10^{-39}	3.29×10^{-44}	1.71×10^{-49}	2.61×10^{-48}
	Intercept	119,335.88	81,652.36	74,442.07	25,967.88	−109,065.03	−100,210.31
	Coeff. 1	1582.96	1328.28	1244.25	1123.73	895.11	910.88
	Coeff. 2		587.88	19.31	22,463.20	15,465.48	16,185.66
	Coeff. 3					424.05	453.40
	Coeff. 4						−2410.34
	Coeff. 5						
	Coeff. 6						
	t_{95}	6.314	2.920	2.920	2.920	2.353	2.132
	t-test_1	21.013	5.790	6.781	10.295	8.713	8.522
	t-test_2		1.175	2.019	5.344	4.015	3.972
	t-test_3					5.918	5.086
	t-test_4						−0.557
	t-test_5						
	t-test_6						
	Significant	Yes	No	No	Yes	Yes	No
	Sign	Yes	Yes	Yes	Yes	Yes	No
	Valid	Yes	No	No	Yes	Yes	No
Order	**Variable\Model n.**	**7**	**8**	**9**	**10**	**11**	**12**
1	ace_i	x	x	x	x	x	x
2	ree_i						
3	ret_i						
4	hsr_i	x	x	x	x	x	x
5	tch_i	x	x	x	x	x	x
6	scs_i						
7	mus_i	x	x	x	x	x	x
8	$d^{AF}{}_i$		x				
9	A_i			x			
10	cap_i				x		
11	$d^{AI}{}_i$					x	
12	$t^{tot}{}_i$						x
13	$d^{tot}{}_i$						
14	huc_i						
	df	4	5	5	5	5	5
	R^2	0.904	0.905	0.904	0.905	0.904	0.904
	$R^2{}_{adj}$	0.901	0.900	0.900	0.900	0.900	0.900
	Significance F	4.40×10^{-53}	6.27×10^{-52}	7.34×10^{-52}	6.67×10^{-52}	7.41×10^{-52}	7.26×10^{-52}

Table 11. *Cont.*

Order	Variable\Model n.	7	8	9	10	11	12
	Intercept	−70,162.32	−48,702.83	−50,856.26	−67,253.63	−69,240.09	−74,550.90
	Coeff. 1	643.63	676.99	646.74	636.86	642.07	643.10
	Coeff. 2	19,847.82	19,229.64	19,966.26	20,251.14	19,896.88	19,886.56
	Coeff. 3	255.53	258.35	257.74	262.26	257.34	257.22
	Coeff. 4	1259.62	1290.13	1250.42	1276.81	1261.47	1257.56
	Coeff. 5		−82,544.79	−1437.26	−53,521.53	−596.02	79,818.04
	Coeff. 6						
	t_{95}	2.132	2.015	2.015	2.015	2.015	2.015
	t-test_1	6.090	5.620	5.985	5.948	5.940	6.056
	t-test_2	5.526	5.120	5.413	5.462	5.431	5.505
	t-test_3	3.497	3.517	3.446	3.509	3.340	3.484
	t-test_4	4.987	4.987	4.798	4.984	4.949	4.953
	t-test_5		−0.583	−0.155	−0.467	−0.078	0.216
	t-test_6						
	Significant	Yes	No	No	No	No	No
	Sign	Yes	No	No	No	No	Yes
	Valid	Yes	No	No	No	No	No
Order	**Variable\Model n.**	**13**	**14**	**15**	**16**	**17**	**18**
1	ace_i	x	x	x	x	x	x
2	ree_i			x			
3	ret_i				x	x	x
4	hsr_i	x	x	x	x	x	x
5	tch_i	x	x	x	x	x	x
6	scs_i					x	
7	mus_i	x	x	x	x	x	x
8	${d^{AF}}_i$						x
9	A_i						
10	cap_i						
11	${d^{AI}}_i$						
12	${t^{tot}}_i$						
13	${d^{tot}}_i$	x					
14	huc_i		x				
	df	5	5	5	5	6	6
	R^2	0.904	0.905	0.908	0.909	0.909	0.909
	${R^2}_{adj}$	0.900	0.901	0.903	0.905	0.904	0.904
	Significance F	7.26×10^{-52}	5.00×10^{-52}	1.10×10^{-52}	6.06×10^{-53}	8.32×10^{-52}	7.62×10^{-52}
	Intercept	−74,550.90	−117,824.59	−111,915.76	−102,072.43	−95,802.97	−133,145.62
	Coeff. 1	643.10	655.79	335.07	385.52	397.19	303.96
	Coeff. 2	19,886.56	20,012.16	714.33	16.23	16.29	18.82
	Coeff. 3	257.22	230.40	18,753.19	17,327.57	17,833.27	17,674.16
	Coeff. 4	1257.56	1296.14	255.40	266.53	288.79	264.87
	Coeff. 5	79,818.04	5288.22	1347.49	1322.12	−1774.62	1295.14
	Coeff. 6					1317.71	99,949.32
	t_{95}	2.015	2.015	2.015	2.015	1.943	1.943
	t-test_1	6.056	6.148	1.786	2.505	2.537	1.502
	t-test_2	5.505	5.559	1.978	2.269	2.268	2.270
	t-test_3	3.484	2.940	5.229	4.690	4.612	4.717
	t-test_4	4.953	5.061	3.543	3.710	3.328	3.673
	t-test_5	0.216	0.892	5.324	5.303	−0.461	5.104
	t-test_6					5.262	0.623
	Significant	No	No	No	Yes	No	No
	Sign	Yes	Yes	Yes	Yes	No	Yes
	Valid	No	No	No	Yes	No	No

Table 12. Models to estimate presences.

Order	Variable\Model n.	1	2	3	4	5	6
1	ace_i	x	x	x	x	x	x
2	ree_i		x				
3	scs_i			x	x	x	x
4	ret_i				x		
5	tch_i					x	
6	mus_i						x
7	hsr_i						
8	${d^{AF}}_i$						
9	A_i						
10	cap_i						
11	${d^{AI}}_i$						
12	${t^{tot}}_i$						
13	${d^{tot}}_i$						
14	huc_i						
	df	1	2	2	3	3	3
	R^2	0.800	0.802	0.832	0.834	0.856	0.868
	${R^2}_{adj}$	0.798	0.798	0.829	0.829	0.852	0.864
	Significance F	6.55×10^{-40}	1.06×10^{-38}	1.36×10^{-42}	1.46×10^{-41}	6.33×10^{-45}	6.73×10^{-47}
	Intercept	251819.80	164437.06	−194,347.93	−142,672.93	−311,053.50	−131,440.04
	Coeff. 1	4320.20	3729.62	3239.32	3670.73	3025.87	2491.73
	Coeff. 2		1363.22	46,339.90	47,182.15	12,875.49	37,193.78
	Coeff. 3				−25.72	1134.04	3909.62
	Coeff. 4						
	Coeff. 5						
	t_{95}	6.314	2.920	2.920	2.353	2.353	2.353
	t-test_1	20.891	5.911	10.634	7.162	10.520	8.175
	t-test_2		0.991	4.545	4.615	1.044	4.025
	t-test_3				−1.046	4.238	5.390
	t-test_4						
	t-test_5						
	Significant	Yes	No	Yes	No	No	Yes
	Sign	Yes	Yes	Yes	No	Yes	Yes
	Valid	Yes	No	Yes	No	No	Yes
Order	**Variable\Model n.**	**7**	**8**	**9**	**10**	**11**	**12**
1	ace_i	x	x	x	x	x	x
2	ree_i						
3	scs_i	x	x	x	x	x	x
4	ret_i						
5	tch_i						
6	mus_i	x	x	x	x	x	x
7	hsr_i	x	x	x	x	x	x
8	${d^{AF}}_i$		x				
9	A_i			x			
10	cap_i				x		
11	${d^{AI}}_i$					x	
12	${t^{tot}}_i$						x
13	${d^{tot}}_i$						
14	huc_i						
	df	4	5	5	5	5	5
	R^2	0.880	0.883	0.880	0.882	0.880	0.881
	${R^2}_{adj}$	0.875	0.877	0.874	0.876	0.874	0.875
	Significance F	7.43×10^{-48}	3.00×10^{-47}	1.11×10^{-46}	4.46×10^{-47}	1.11×10^{-46}	6.83×10^{-47}
	Intercept	−147,521.11	−320,838.73	−142,610.04	−136,636.34	−148,655.35	−559,850.42
	Coeff. 1	1969.65	1701.92	1970.38	1874.88	1971.85	1989.77
	Coeff. 2	22,547.13	20,257.72	22,560.66	27,215.72	22,472.83	21,104.26
	Coeff. 3	4334.67	4076.01	4333.01	4509.78	4330.52	4465.27
	Coeff. 4	37,138.76	43,157.54	37,171.69	39,546.60	37,082.77	35,183.38
	Coeff. 5		703,651.96	−360.66	−491,020.93	683.10	7,719,931.11
	t_{95}	2.132	2.015	2.015	2.015	2.015	2.015
	t-test_1	5.909	4.608	5.802	5.526	5.748	5.958
	t-test_2	2.269	2.034	2.247	2.597	2.183	2.101
	t-test_3	6.130	5.665	6.000	6.298	5.979	6.206
	t-test_4	3.249	3.618	3.159	3.432	3.186	3.033
	t-test_5		1.629	−0.013	−1.357	0.030	0.986

Table 12. *Cont.*

Order	**Variable\Model n.**	**7**	**8**	**9**	**10**	**11**	**12**	
	Significant	Yes	No	No	No	No	No	
	Sign	Yes	Yes	No	No	Yes	Yes	
	Valid	Yes	No	No	No	No	No	
Order	**Variable\Model n.**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
1	ace_i	x	x	x	x	x	x	x
2	ree_i			x				
3	scs_i	x	x	x	x	x		
4	ret_i				x			x
5	tch_i					x	x	x
6	mus_i	x	x	x	x	x	x	x
7	hsr_i	x	x	x	x	x	x	x
8	${d^{AF}}_i$							
9	A_i							
10	cap_i							
11	${d^{AI}}_i$							
12	${t^{tot}}_i$							
13	${d^{tot}}_i$	x						
14	huc_i		x					
	df	5	5	5	5	5	4	5
	R^2	0.881	0.880	0.884	0.881	0.886	0.885	0.886
	${R^2}_{adj}$	0.875	0.875	0.878	0.876	0.880	0.881	0.881
	Significance F	7.53×10^{-47}	9.76×10^{-47}	1.92×10^{-47}	6.22×10^{-47}	9.17×10^{-48}	7.48×10^{-49}	7.34×10^{-48}
	Intercept	−509,712.73	−228,911.02	−268,157.62	−103,344.33	−218,173.27	−190,198.61	−149,673.90
	Coeff. 1	1991.25	1999.84	1066.82	2351.66	2030.69	2085.68	2413.48
	Coeff. 2	21,403.22	20,872.86	2095.06	22,291.56	7785.26	695.11	−20.61
	Coeff. 3	4424.06	4357.63	22,283.38	−23.99	597.61	3616.30	681.14
	Coeff. 4	35,462.46	37,500.76	4596.37	4220.73	3636.55	36,022.00	3536.93
	Coeff. 5	6,198,042.71	8976.50	34,058.84	40,696.45	33,766.31		39,222.63
	t_{95}	2.015	2.015	2.015	2.015	2.015	2.132	2.015
	t-test_1	5.952	5.884	1.840	4.835	6.189	6.587	5.132
	t-test_2	2.134	1.985	1.891	2.245	0.664	3.175	−0.943
	t-test_3	6.186	6.127	2.270	−1.078	2.263	4.779	3.103
	t-test_4	3.057	3.263	6.453	5.907	4.789	3.347	4.643
	t-test_5	0.879	0.501	2.985	3.423	2.985		3.474
	Significant	No	No	No	No	No	Yes	No
	Sign	Yes	Yes	Yes	No	Yes	Yes	No
	Valid	No	No	No	No	No	Yes	No

Overall, 18 models for estimating arrivals and 19 models for estimating presences were calibrated; of these models, five models for estimating arrivals and five models for estimating presences were valid in terms of significance and sign of the coefficients. At the end of the procedure, model no. 16 for estimating arrivals and model no. 18 for estimating presences were identified as the best. These models have the maximum values of R^2 and ${R^2}_{adj}$, and comply with all the significance tests. The values of the coefficients of determination are sufficiently high in both cases (0.909 for arrivals and 0.885 for presences). It is important to note that in both models, the only accessibility variable found to be significant is the one related to high-speed rail services, hsr_i. The other accessibility variables were not statistically significant.

Figures 4 and 5 show the scatter diagrams comparing the actual and estimated values.

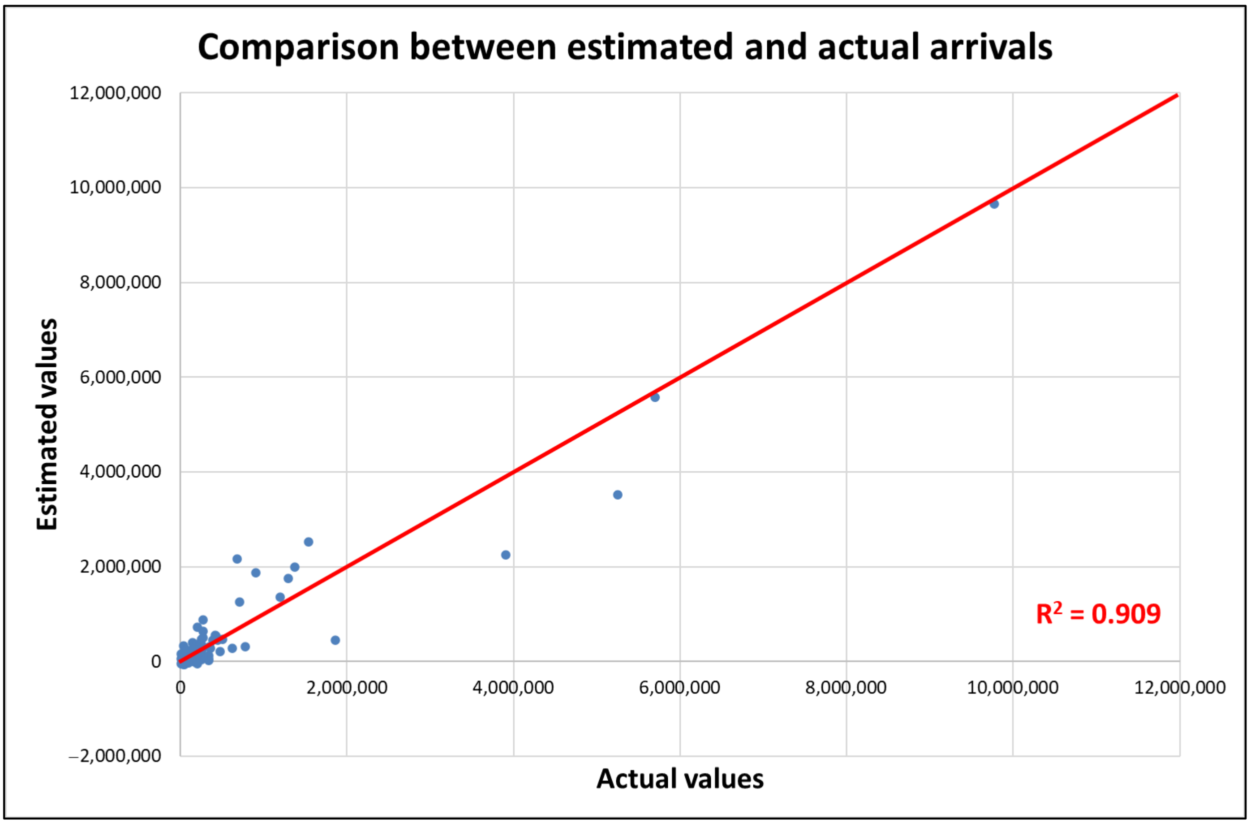


Figure 4. Model for estimating arrivals: comparison between real and model data.

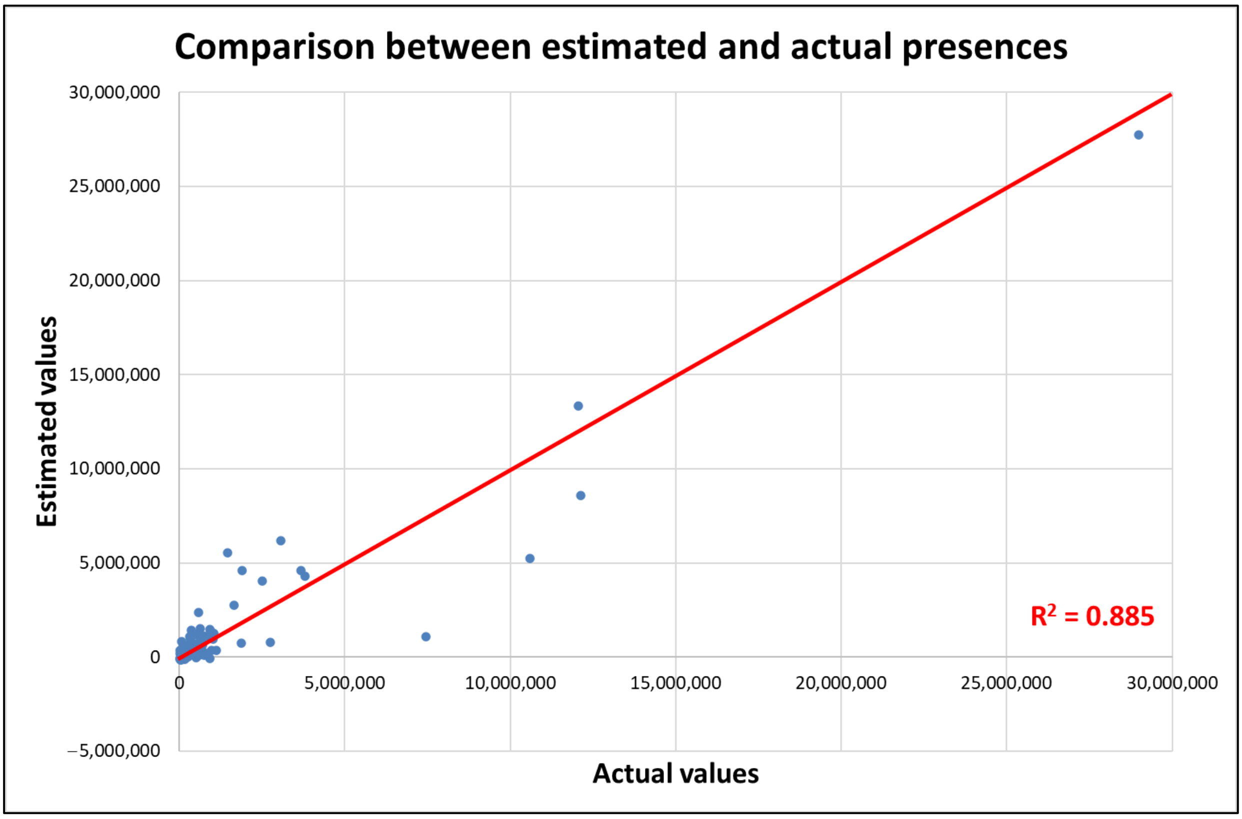


Figure 5. Model for estimating presences: comparison between real and model data.

The best models are formulated as follows:

$$ARR_i = -102{,}072.43 + 385.52\ ace_i + 16.23\ ret_i + 17{,}327.57\ hsr_i + 266.53\ tch_i + 1322.12\ mus_i \quad (1)$$

$$PRE_i = -190{,}198.61 + 2085.68\ ace_i + 695.11\ tch_i + 3616.30\ mus_i + 36{,}022.00\ hsr_i \quad (2)$$

The analysis of these models highlights the following aspects:

(a) The intercept assumes a negative value. This property permits the models to be used only for overall evaluations of the entire set of municipalities (remember that, having used the generalised least squares method, the sum of the values estimated by the model for all the municipalities is equal to the sum of the true values). The application to a specific municipality could give implausible values, and for municipalities with less tourist importance, negative values.

(b) The variables linked to creative, artistic, and entertainment activities, total cultural assets, the presence of libraries, museums and other cultural activities, and direct rides on high-speed services always appear in both models. In the arrivals model, the variable related to commercial activities is also significant, while it is not statistically significant for presences. This indicates that commercial activities have a greater influence on shorter-duration trips than longer ones. In all cases, the variables closely linked to the tourist offer of the place of destination are significant.

(c) Among the accessibility variables, only the one representing high-speed rail services is statistically significant in estimating arrivals and presences. The other accessibility variables, at least for the provincial capitals, are not influential.

To evaluate the importance of high-speed services with respect to the other factors, a sensitivity analysis was carried out, increasing the overall values of each variable by 10% and evaluating the percentage increase in the number of arrivals and presences. The results are summarised in Tables 13 and 14.

Table 13. Results of the sensitivity analysis for arrivals.

Total Arrivals	Variable	Total Arrivals by Model (+10% of Variable)	Percentage Variation
48,758,419	ace_i	49,623,298	+1.77%
	ret_i	49,918,715	+2.17%
	hsr_i	50,352,555	+3.27%
	tch_i	50,427,540	+3.42%
	mus_i	49,581,833	+1,69%

Table 14. Results of the sensitivity analysis for presences.

Total Presences	Variable	Total Presences by Model (+10% of Variable)	Percentage Variation
124,871,320	ace_i	129,550,341	+3.75%
	tch_i	129,224,383	+3.49%
	mus_i	127,123,549	+1.80%
	hsr_i	128,185,344	+2.65%

The analysis of these results leads to the following considerations:

- High-speed rail services have an important impact on the flow of arrivals and presences in accommodation facilities. The elasticity is greater for arrivals, where an increase of +10% in supply can be estimated as a +3.27% increase in arrivals, while this value is reduced to +2.65% for presences. In both cases, the values are significant: for arrivals, the elasticity is second only to that linked to total cultural assets, while for presences, it is third, being also preceded by creative, artistic, and entertainment activities.

- A comparison of the model's elasticities between arrivals and presences shows that there is practically the same elasticity for the variable on total cultural heritage (+3.42% arrivals and +3.49% presences), highlighting how this explanatory variable has more or less the same effect on all stays, regardless of their duration. On the other hand, creative, artistic, and entertainment activities have a greater elasticity on arrivals than on presences, showing a tendency to influence shorter stays more.
- Museums, libraries, and other cultural activities have practically the same elasticity, as total cultural heritage, on both arrivals and presences (+1.69% arrivals and +1.80% presences).
- Commercial activities, as already mentioned, show an influence only on arrivals and, therefore, a greater influence on shorter stays.

From the calibration of these models and analyses, it can be concluded that the impact of high-speed rail services on tourism flows, as measured by arrivals and presences in accommodation establishments, is significant. For arrivals, the elasticity of the variable is high, of the same order of magnitude as for the total number of cultural assets. For presences, it is lower, but still very significant. Another fact to note is that, of all the accessibility variables considered, high-speed services are the only statistically significant.

5. An Application to a Case Study

The calibrated models were applied to a specific case study, the city of Benevento. Benevento is a small-medium-sized provincial capital with about 60,000 inhabitants (only 36 out of 111 provincial capitals have fewer residents than Benevento), but it has several important historical/archaeological sites, including the monumental complex of Santa Sofia, a UNESCO World Heritage Site, the Arch of Trajan, the Roman Theatre and the Rocca dei Rettori, as well as several museums and churches of great value. The accessibility of the city, however, is not as good as the artistic and historical heritage: the railway connections with the regional capital (Naples) are not efficient and have a modest frequency, while Trenitalia's Frecce services connect the city on the Rome-Bari route, for a total of only 28 runs, as the sum of those arriving and departing.

Currently (2018 data, pre-COVID), annual arrivals in accommodation amount to 36,252 (on average, 99 per day), while presences stand at 80,144 (on average, 220 per day), with an accommodation supply of 57 establishments for a total of 1039 beds; therefore, there are about 35 arrivals and 77 presences per bed; for comparison, the city of Naples has about 87 arrivals and 232 presences per bed.

A new HS railway line is currently under construction, which will serve a Naples-Benevento-Bari route, with a maximum line speed of 250 km/h. As it is under construction, the frequency of services has not yet been established, but it can be assumed that the service will be organised on 12 pairs of daily runs, increasing the service to a total of 40 daily runs, as the sum of arriving and departing runs.

Assuming the same elasticity as estimated in Section 4, the new services would increase the current services by 42.8% and, therefore, could lead to an increase, other factors unchanged, of 15.8% in arrivals (+5728) and 11.3% in presences (+9056), increasing the overall annual occupancy rate from 21.1% to 23.5%.

These results further underline how high-speed railway lines can have a significant impact on tourist attractiveness.

6. Conclusions

High-speed rail transport systems have proved to be an important tool for spatial development all over the world. Increased accessibility due to HSR services has a positive impact on industrial and commercial activities, increases property values, and reduces emissions of pollutants and greenhouse gases. This paper studied the effects on tourism in Italy, taking provincial capitals as territorial reference units. The calibrated multiple linear regression models showed that HSR services are the most important accessibility factor for tourism and that their impact is significant in terms of arrivals and presence in accommodation facilities.

We can summarise the results of this study as follows: (a) among all the accessibility variables, the availability of HSR services has a significant impact on tourism attractiveness in Italy; (b) the elasticity analysis showed that the influence of this variable is of the same magnitude as the variables related to the offer of historical–cultural assets and creative, artistic and entertainment activities; (c) the application to the case study of Benevento showed that the presence of HSR services is fundamental for the city's tourism development.

Based on these results, we can reasonably affirm that HSR services represent a strategic factor for the development and promotion of tourism in Italy. With an equal historical and cultural offer, the locations better served by rail transport show a higher attractiveness. This factor should be included in the assessments of policymakers when investing in HSR systems. Indeed, from a practical point of view, the possibility of forecasting the impact on tourism of HSR services (and of the infrastructures they use) makes it possible to make more conscious decisions in the choice of investments, being then able to derive from the results also obtained the impacts on the socio-economic development of the territories involved.

In this regard, further developments of this research will be aimed at extending the study to municipalities that are not provincial capitals and at considering multimodal transport accessibility variables, directly considering the interchanges between different transport systems and their role in improving the sustainable way of enjoying territory.

Author Contributions: Conceptualization, M.G. and R.A.L.R.; methodology, M.G. and R.A.L.R.; model, M.G.; numerical results, M.G. and R.A.L.R.; resources, M.G.; data curation, M.G. and R.A.L.R.; writing—original draft preparation, M.G. and R.A.L.R.; writing—review and editing, M.G. and R.A.L.R. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All available data are quoted in the References or available in Appendix A.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

This appendix reports the input data used in the paper.

Table A1. Variables on the supply of historical/cultural assets.

City	scs_i	tch_i	mus_i	huc_i	City	scs_i	tch_i	mus_i	huc_i
Agrigento	7	151	17	5.96	Mantova	14	355	49	22.38
Alessandria	9	135	11	11.37	Massa	3	169	0	9.53
Ancona	13	663	4	9.62	Matera	5	210	9	5.53
Andria	1	90	0	5.55	Messina	8	396	4	7.51
Aosta	2	56	1	8.77	Milan	61	1691	271	8.49
Arezzo	16	520	1	10.39	Modena	29	467	12	11.26
Ascoli Pic.	19	943	46	17.02	Monza	2	127	6	7.43
Asti	9	174	4	14.06	Naples	56	1528	34	20.38
Avellino	5	97	6	1.66	Novara	7	275	1	3.29
Bari	16	641	3	6.32	Nuoro	5	67	51	1.87
Barletta	5	184	1	17.59	Oristano	3	69	39	3.57
Belluno	2	267	2	15.31	Padua	10	1295	1	5.63
Benevento	8	323	0	2.99	Palermo	32	799	1223	5.95
Bergamo	11	653	18	12.69	Parma	27	802	3	12.02
Biella	6	255	4	16.15	Pavia	13	477	5	14.12
Bologna	73	1791	170	12.76	Perugia	33	735	208	9.69

Table A1. *Cont.*

City	scs_i	tch_i	mus_i	huc_i	City	scs_i	tch_i	mus_i	huc_i
Bolzano	9	97	9	7.98	Pesaro	10	479	1	7.40
Brescia	18	794	3	15.04	Pescara	7	118	5	2.61
Brindisi	5	107	3	4.30	Piacenza	14	652	1	8.36
Cagliari	29	541	96	5.70	Pisa	20	702	27	17.26
Caltanissetta	3	124	0	7.97	Prato	13	210	10	6.74
Campobasso	4	214	6	10.59	Ragusa	2	165	1	5.16
Carbonia	4	270	48	1.22	Ravenna	22	708	7	5.04
Caserta	7	150	0	10.70	Reggio Cal.	20	336	3	0.81
Catania	9	396	7	10.03	Reggio Emil.	13	347	4	7.75
Catanzaro	12	191	0	3.17	Rieti	2	175	0	15.13
Cesena	12	300	1	4.47	Rimini	9	363	4	1.29
Chieti	7	176	0	5.68	Rome	153	6239	1946	3.80
Como	8	300	5	12.13	Rovigo	2	250	13	5.26
Cosenza	7	177	1	11.62	Salerno	14	310	5	8.10
Cremona	8	211	8	18.07	Sassari	7	313	58	8.26
Crotone	6	47	0	1.77	Savona	9	675	0	16.17
Cuneo	8	190	2	11.91	Siena	42	1675	47	27.05
Enna	1	80	1	6.48	Siracusa	11	254	4	3.74
Fermo	5	423	0	11.56	Sondrio	2	75	9	14.34
Ferrara	23	1699	1	13.19	Taranto	5	390	2	2.85
Florence	104	2093	105	23.98	Teramo	6	176	0	17.77
Foggia	7	275	2	1.77	Terni	11	196	13	6.43
Forlì	10	436	6	8.28	Turin	57	1729	53	8.57
Frosinone	1	30	1	1.68	Trani	4	100	0	18.48
Genoa	58	4356	157	20.23	Trapani	1	103	0	7.67
Gorizia	3	356	0	14.87	Trento	2	139	23	12.97
Grosseto	5	322	3	4.17	Treviso	2	876	23	4.67
Imperia	9	365	1	18.42	Trieste	38	1496	20	24.34
Isernia	5	141	0	22.11	Udine	5	474	71	6.72
La Spezia	9	344	24	15.77	Urbino	8	504	0	29.74
L'Aquila	7	690	0	14.04	Varese	6	157	1	11.45
Latina	9	46	3	0.99	Venice	26	3790	1031	20.59
Lecce	10	584	1	5.39	Verbania	1	139	0	22.74
Lecco	7	103	0	20.12	Vercelli	3	199	0	15.51
Livorno	8	283	83	22.23	Verona	22	1711	6	11.48
Lodi	6	165	3	9.85	Vibo Valentia	3	147	3	3.58
Lucca	11	438	11	30.56	Vicenza	8	880	6	9.68
Macerata	4	370	1	12.11	Viterbo	14	487	0	14.35

Table A2. Entertainment/amusement variables.

City	ace_i	ree_i	City	ace_i	ree_i	City	ace_i	ree_i
Agrigento	18	34	Foggia	30	47	Pistoia	51	23
Alessandria	89	63	Forlì	59	86	Pordenone	111	114
Ancona	73	101	Frosinone	13	40	Potenza	14	87
Andria	14	14	Genoa	425	357	Prato	108	90
Aosta	71	11	Gorizia	42	16	Ragusa	28	56
Arezzo	99	75	Grosseto	47	138	Ravenna	173	331
Ascoli Piceno	36	25	Imperia	19	52	Reggio Cal.	17	81
Asti	25	91	Isernia	4	14	Reggio Emil.	137	133
Avellino	20	50	La Spezia	32	47	Rieti	15	21
Bari	131	83	L'Aquila	52	6	Rimini	193	717
Barletta	14	31	Latina	59	168	Rome	6171	2429
Belluno	24	5	Lecce	45	67	Rovigo	44	39
Benevento	24	53	Lecco	27	38	Salerno	120	167
Bergamo	175	138	Livorno	106	133	Sassari	136	95
Biella	17	31	Lodi	26	14	Savona	85	71
Bologna	700	200	Lucca	124	106	Siena	48	22
Bolzano	106	60	Macerata	33	30	Siracusa	49	68
Brescia	486	286	Mantova	48	43	Sondrio	10	24
Brindisi	17	25	Massa	28	212	Taranto	22	40
Cagliari	97	208	Matera	21	30	Teramo	23	36
Caltanissetta	10	34	Messina	47	109	Terni	41	105
Campobasso	16	46	Milan	3828	2098	Turin	941	634
Carbonia	3	3	Modena	266	396	Trani	10	19
Caserta	45	65	Monza	144	39	Trapani	7	37
Catania	118	271	Naples	890	797	Trento	371	59
Catanzaro	14	81	Novara	64	67	Treviso	105	132
Cesena	146	100	Nuoro	16	6	Trieste	185	59
Chieti	19	14	Oristano	33	34	Udine	84	338
Como	52	48	Padua	209	108	Urbino	38	4
Cosenza	26	24	Palermo	370	302	Varese	66	44
Cremona	104	81	Parma	179	166	Venice	368	203
Crotone	50	27	Pavia	56	40	Verbania	10	16
Cuneo	21	82	Perugia	133	177	Vercelli	21	48
Enna	15	9	Pesaro	83	105	Verona	887	394
Fermo	27	128	Pescara	93	177	Vibo Valentia	1	20
Ferrara	67	171	Piacenza	224	120	Vicenza	93	290
Florence	722	356	Pisa	53	185	Viterbo	32	94

Table A3. Retail variables.

City	ret_i	City	ret_i	City	ret_i
Agrigento	1795	Foggia	1756	Pistoia	2802
Alessandria	2257	Forlì	5345	Pordenone	1639
Ancona	2854	Frosinone	2289	Potenza	2207
Andria	1976	Genoa	18762	Prato	9081
Aosta	1537	Gorizia	940	Ragusa	2638
Arezzo	4333	Grosseto	3620	Ravenna	3970
Ascoli Piceno	3114	Imperia	1344	Reggio Cal.	4923
Asti	2377	Isernia	693	Reggio Emil.	9646
Avellino	1927	La Spezia	2812	Rieti	1349
Bari	4128	L'Aquila	1985	Rimini	5328
Barletta	1230	Latina	4556	Rome	83,495
Belluno	1490	Lecce	1604	Rovigo	1362
Benevento	2028	Lecco	1535	Salerno	4411
Bergamo	6795	Livorno	4573	Sassari	4354
Biella	1880	Lodi	1056	Savona	2355
Bologna	11,243	Lucca	2642	Siena	2200
Bolzano	4162	Macerata	1432	Siracusa	2922
Brescia	7324	Mantova	1556	Sondrio	781
Brindisi	902	Massa	2310	Taranto	2105
Cagliari	6496	Matera	1751	Teramo	1477
Caltanissetta	2117	Messina	5859	Terni	4633
Campobasso	1450	Milan	108,407	Turin	26,629
Carbonia	991	Modena	10,935	Trani	692
Caserta	2350	Monza	4136	Trapani	2394
Catania	9689	Naples	28621	Trento	5300
Catanzaro	3492	Novara	3400	Treviso	2594
Cesena	3101	Nuoro	1246	Trieste	5640
Chieti	1892	Oristano	1491	Udine	2403
Como	2722	Padua	9394	Urbino	493
Cosenza	2490	Palermo	18,327	Varese	2962
Cremona	1727	Parma	5244	Venice	19,250
Crotone	1839	Pavia	2125	Verbania	827
Cuneo	1931	Perugia	5829	Vercelli	1376
Enna	703	Pesaro	2693	Verona	5971
Fermo	945	Pescara	4581	Vibo Valentia	1319
Ferrara	3066	Piacenza	3009	Vicenza	3409
Florence	22,876	Pisa	2716	Viterbo	2685

Table A4. Accessibility vriables.

City	hsr_i	$d^{AF}{}_i$	$d^{AI}{}_i$	A_i	$t^{tot}{}_i$	$d^{tot}{}_i$
Agrigento	0	0.1427	0.647	6.708	41.677	32.300
Alessandria	1	0.1886	1.427	18.272	15.121	14.085
Ancona	14	0.3576	5.358	14.340	15.074	13.718
Andria	0	0.2494	1.723	11.692	22.299	21.204
Aosta	0	0.1428	0.883	11.150	19.320	18.168
Arezzo	3	0.4531	1.280	16.152	13.941	12.918
Ascoli Piceno	3	0.4863	1.202	13.213	16.519	15.260
Asti	1	0.1756	1.593	16.798	16.155	15.028
Avellino	0	0.3699	1.801	14.703	20.295	19.665
Bari	15	0.2234	20.000	13.844	22.939	22.390
Barletta	13	0.2544	1.564	12.409	21.674	20.991
Belluno	0	0.1706	0.956	13.211	16.668	14.828
Benevento	5	0.3992	1.469	13.550	20.139	19.042
Bergamo	2	0.1718	1.173	22.416	14.631	13.539
Biella	0	0.1580	1.403	14.612	17.169	15.912
Bologna	85	0.2668	20.000	22.396	12.630	11.659
Bolzano	5	0.1567	20.000	13.157	16.623	15.474
Brescia	27	0.1877	0.776	20.299	13.885	12.728
Brindisi	9	0.1780	0.834	10.062	26.496	25.671
Cagliari	0	0.1918	13.451	9.764	25.663	28.142
Caltanissetta	0	0.1423	0.926	7.349	40.014	32.191
Campobasso	10	0.3997	0.826	11.816	19.761	18.410
Carbonia	0	0.1846	1.860	7.506	28.080	29.056
Caserta	5	0.4516	3.081	16.666	19.237	18.703
Catania	0	0.1280	20.000	10.496	36.828	33.881
Catanzaro	4	0.1661	2.787	8.220	29.568	28.463
Cesena	0	0.3054	1.140	16.872	13.733	12.251
Chieti	0	0.4436	7.642	14.318	17.023	15.978
Como	0	0.1626	2.354	20.684	15.540	14.424
Cosenza	5	0.1849	1.451	8.913	27.196	26.664
Cremona	0	0.1978	0.786	19.034	13.687	12.463
Crotone	0	0.1637	0.971	7.685	30.021	28.434
Cuneo	0	0.1679	1.030	12.930	17.985	16.793
Enna	0	0.1405	1.170	7.175	39.317	32.371
Fermo	0	0.3849	1.308	13.518	15.813	14.489
Ferrara	2	0.2355	1.984	18.313	13.515	11.989
Florence	58	0.3563	1.186	21.471	13.104	12.202
Foggia	14	0.2935	0.788	12.493	20.396	19.405
Forlì	4	0.2959	1.479	18.046	13.426	12.084
Frosinone	0	0.9188	0.919	16.021	17.225	16.635
Gorizia	0	0.1616	5.181	11.902	18.068	16.676
Grosseto	6	0.6416	0.650	13.271	15.454	13.933
Imperia	0	0.1724	0.850	11.239	18.547	17.356
Isernia	0	0.4931	0.972	12.639	19.066	17.847
La Spezia	7	0.2659	1.257	15.732	14.571	13.095
L'Aquila	0	0.6504	1.047	14.021	17.005	16.061
Latina	0	1.2230	1.223	15.507	18.161	16.821
Lecce	9	0.1676	0.644	9.263	27.828	26.747
Lecco	0	0.1633	1.416	19.155	15.513	14.295
Livorno	6	0.3515	3.998	15.790	14.848	13.227

Table A4. *Cont.*

City	hsr_i	${d^{AF}}_i$	${d^{AI}}_i$	A_i	${t^{tot}}_i$	${d^{tot}}_i$
Lodi	0	0.1840	1.290	21.598	14.150	13.040
Lucca	0	0.3093	3.296	16.914	14.203	12.755
Macerata	1	0.4147	1.851	13.550	15.876	14.092
Mantova	1	0.2138	0.979	19.896	13.096	12.058
Massa	6	0.2858	1.836	15.874	14.556	13.024
Matera	1	0.2249	1.653	11.165	23.997	22.640
Messina	0	0.1445	0.962	9.051	33.849	31.025
Milan	94	0.1744	2.301	30.949	14.458	13.510
Modena	11	0.2478	2.573	21.135	12.801	11.719
Monza	0	0.1708	2.103	28.530	14.595	13.702
Naples	48	0.4210	20.000	21.502	19.725	19.165
Novara	0	0.1694	3.475	19.735	15.535	14.451
Nuoro	0	0.2672	0.608	6.287	31.754	25.070
Oristano	0	0.2281	1.113	7.106	29.003	26.603
Padova	45	0.2040	2.421	20.855	13.952	12.635
Palermo	0	0.1745	0.480	10.517	40.653	28.651
Parma	10	0.2199	1.093	19.569	13.262	12.054
Pavia	2	0.1783	1.471	21.159	14.562	13.460
Perugia	1	0.5343	20.000	14.443	15.402	13.556
Pesaro	16	0.3394	1.746	15.317	14.269	12.889
Pescara	12	0.4226	20.000	14.412	16.961	15.782
Piacenza	10	0.1973	0.891	20.347	13.544	12.586
Pisa	7	0.3269	20.000	16.567	14.499	13.038
Pistoia	0	0.3144	1.367	18.365	13.594	12.440
Pordenone	3	0.1780	1.360	14.892	16.481	14.687
Potenza	1	0.2602	0.810	11.389	22.269	21.404
Prato	0	0.3386	1.237	20.876	13.231	12.275
Ragusa	0	0.1237	1.105	7.069	40.111	35.135
Ravenna	3	0.2797	1.289	17.386	13.653	12.136
Reggio di Calabria	4	0.1431	0.756	8.144	32.242	31.266
Reggio nell'Emilia	30	0.2336	1.566	20.227	13.040	11.883
Rieti	0	0.9465	0.946	14.772	16.487	14.975
Messina	0	0.1445	0.962	9.051	33.849	31.025
Milan	94	0.1744	2.301	30.949	14.458	13.510
Modena	11	0.2478	2.573	21.135	12.801	11.719
Monza	0	0.1708	2.103	28.530	14.595	13.702
Naples	48	0.4210	20.000	21.502	19.725	19.165
Novara	0	0.1694	3.475	19.735	15.535	14.451
Nuoro	0	0.2672	0.608	6.287	31.754	25.070
Oristano	0	0.2281	1.113	7.106	29.003	26.603
Padova	45	0.2040	2.421	20.855	13.952	12.635
Palermo	0	0.1745	0.480	10.517	40.653	28.651
Parma	10	0.2199	1.093	19.569	13.262	12.054
Pavia	2	0.1783	1.471	21.159	14.562	13.460
Perugia	1	0.5343	20.000	14.443	15.402	13.556
Pesaro	16	0.3394	1.746	15.317	14.269	12.889
Pescara	12	0.4226	20.000	14.412	16.961	15.782
Piacenza	10	0.1973	0.891	20.347	13.544	12.586
Pisa	7	0.3269	20.000	16.567	14.499	13.038
Pistoia	0	0.3144	1.367	18.365	13.594	12.440
Pordenone	3	0.1780	1.360	14.892	16.481	14.687

Table A4. *Cont.*

City	hsr_i	$d^{AF}{}_i$	$d^{AI}{}_i$	A_i	$t^{tot}{}_i$	$d^{tot}{}_i$
Potenza	1	0.2602	0.810	11.389	22.269	21.404
Prato	0	0.3386	1.237	20.876	13.231	12.275
Ragusa	0	0.1237	1.105	7.069	40.111	35.135
Ravenna	3	0.2797	1.289	17.386	13.653	12.136
Reggio di Calabria	4	0.1431	0.756	8.144	32.242	31.266
Reggio nell'Emilia	30	0.2336	1.566	20.227	13.040	11.883
Rieti	0	0.9465	0.946	14.772	16.487	14.975
Rimini	17	0.3110	1.091	16.367	13.935	12.544
Rome	103	3.7307	20.000	30.293	16.718	15.565
Rovigo	2	0.2204	1.257	18.089	13.725	12.247
Salerno	15	0.3453	1.593	14.849	20.746	20.267
Sassari	0	0.2495	0.461	6.511	33.872	25.953
Savona	0	0.1947	1.948	14.117	16.574	15.275
Siena	0	0.4643	0.936	15.673	14.007	12.784
Siracusa	0	0.1199	1.783	7.822	38.608	35.587
Sondrio	0	0.1477	0.691	11.671	17.887	15.928
Taranto	2	0.1943	1.078	11.470	25.302	24.511
Teramo	1	0.4916	1.651	13.636	16.637	15.556
Terni	3	0.8848	1.084	15.232	15.880	14.487
Turin	31	0.1614	6.587	21.782	17.123	16.159
Trani	0	0.2465	2.140	12.459	22.123	21.297
Trapani	0	0.1527	0.352	6.555	43.488	31.255
Trento	5	0.1720	1.766	15.373	15.148	13.994
Treviso	3	0.1942	2.996	18.666	15.034	13.354
Trieste	4	0.1559	3.025	11.761	18.777	17.425
Udine	3	0.1655	2.691	12.953	17.726	16.089
Urbino	0	0.3730	1.186	13.751	14.980	13.064
Varese	0	0.1601	3.488	18.817	15.914	14.766
Venice	46	0.1981	20.000	18.758	14.705	13.402
Verbania	0	0.1516	1.650	19.053	17.101	16.007
Vercelli	0	0.1705	1.923	17.808	15.903	14.734
Verona	32	0.2016	0.851	20.637	13.589	12.286
Vibo Valentia	2	0.1660	2.493	8.097	29.479	28.393
Vicenza	24	0.1966	1.417	19.590	14.130	12.656
Viterbo	0	1.2315	1.231	14.566	15.962	14.480

Table A5. Importance variables.

City	cap_i	City	cap_i	City	cap_i	City	cap_i
Agrigento	0	Como	0	Matera	0	Rome	1
Alessandria	0	Cosenza	0	Messina	0	Rovigo	0
Ancona	1	Cremona	0	Milan	1	Salerno	0
Andria	0	Crotone	0	Modena	0	Sassari	0
Aosta	1	Cuneo	0	Monza	0	Savona	0
Arezzo	0	Enna	0	Naples	1	Siena	0

Table A5. *Cont.*

City	cap_i	City	cap_i	City	cap_i	City	cap_i
Ascoli Piceno	0	Fermo	0	Novara	0	Siracusa	0
Asti	0	Ferrara	0	Nuoro	0	Sondrio	0
Avellino	0	Florence	1	Oristano	0	Taranto	0
Bari	1	Foggia	0	Padova	0	Teramo	0
Barletta	0	Forlì	0	Palermo	1	Terni	0
Belluno	0	Frosinone	0	Parma	0	Turin	1
Benevento	0	Genoa	1	Pavia	0	Trani	0
Bergamo	0	Gorizia	0	Perugia	1	Trapani	0
Biella	0	Grosseto	0	Pesaro	0	Trento	1
Bologna	1	Imperia	0	Pescara	0	Treviso	0
Bolzano	0	Isernia	0	Piacenza	0	Trieste	1
Brescia	0	La Spezia	0	Pisa	0	Udine	0
Brindisi	0	L'Aquila	1	Pistoia	0	Urbino	0
Cagliari	1	Latina	0	Pordenone	0	Varese	0
Caltanissetta	0	Lecce	0	Potenza	1	Venezia	1
Campobasso	1	Lecco	0	Prato	0	Verbania	0
Carbonia	0	Livorno	0	Ragusa	0	Vercelli	0
Caserta	0	Lodi	0	Ravenna	0	Verona	0
Catania	0	Lucca	0	Reggio Calabria	0	Vibo Valentia	0
Catanzaro	1	Macerata	0	Reggio Emilia	0	Vicenza	0
Cesena	0	Mantova	0	Rieti	0	Viterbo	0
Chieti	0	Massa	0	Rimini	0		

References

1. UIC. Intercity and High-Speed. 2021. Available online: https://uic.org/passenger/highspeed/ (accessed on 30 September 2022).
2. UNWTO. *UNWTO Tourism Highlights*, 2018 ed.; UNWTO: Madrid, Spain, 2019.
3. UNWTO. Glossary of Tourism Terms. 2022. Available online: https://www.unwto.org/glossary-tourism-terms (accessed on 30 October 2021).
4. United Nations. International Recommendations for Tourism Statistics 2008. Available online: https://unstats.un.org/unsd/publication/seriesm/seriesm_83rev1e.pdf (accessed on 30 October 2021).
5. UNESCO. World Heritage List Statistics. 2022. Available online: https://whc.unesco.org/en/list/stat (accessed on 30 October 2021).
6. Song, H.; Li, G. Tourism demand modelling and forecasting—A review of recent research. *Tour. Manag.* **2008**, *29*, 203–220. [CrossRef]
7. Cho, V. A comparison of three different approaches to tourist arrival forecasting. *Tour. Manag.* **2003**, *24*, 323–330. [CrossRef]
8. Palmer, A.; Montano, J.J.; Sesé, A. Designing an artificial neural network for forecasting tourism time series. *Tour. Manag.* **2006**, *27*, 781–790. [CrossRef]
9. Akin, M. A novel approach to model selection in tourism demand modelling. *Tour. Manag.* **2015**, *48*, 64–72. [CrossRef]
10. Marrocu, E.; Paci, R. Different tourists to different destinations. Evidence from spatial interaction models. *Tour. Manag.* **2013**, *39*, 71–83. [CrossRef]
11. Chan, F.; Lim, C. Spectral analysis of seasonality in tourism demand. *Math. Comput. Simul.* **2011**, *81*, 1409–1418. [CrossRef]
12. Hadavandi, E.; Ghanbari, A.; Shahanaghi, K.; Abbasian-Naghneh, S. Tourist arrival forecasting by evolutionary fuzzy systems. *Tour. Manag.* **2011**, *32*, 1196–1203. [CrossRef]
13. Hassani, H.; Webster, A.; Sirimal Silva, E.; Heravi, S. Forecasting U.S. Tourist arrivals using optimal Singular Spectrum Analysis. *Tour. Manag.* **2015**, *46*, 322–335. [CrossRef]
14. Li, S.; Chen, T.; Wang, L.; Ming, C. Effective tourist volume forecasting supported by PCA and improved BPNN using Baidu index. *Tour. Manag.* **2018**, *68*, 116–126. [CrossRef]
15. Greenidge, K. Forecasting tourism demand. An STM approach. *Ann. Tour. Res.* **2001**, *28*, 98–112. [CrossRef]
16. Song, H.; Li, G.; Witt, S.F.; Athanasopoulos, G. Forecasting tourist arrivals using time-varying parameter structural time series models. *Int. J. Forecast.* **2011**, *27*, 855–869. [CrossRef]
17. Goh, C.; Law, R. Modeling and forecasting tourism demand for arrivals with stochastic nonstationary seasonality and intervention. *Tour. Manag.* **2002**, *23*, 449–510. [CrossRef]
18. Chu, F.-L. A fractionally integrated autoregressive moving average approach to forecasting tourism demand. *Tour. Manag.* **2008**, *29*, 79–88. [CrossRef] [PubMed]

19. Andrawis, R.R.; Atiya, A.F.; El-Shishiny, H. Combination of long term and short term forecasts, with application to tourism demand forecasting. *Int. J. Forecast.* **2011**, *27*, 870–886. [CrossRef]
20. Nelson, L.A.; Dickey, D.A.; Smith, J.M. Estimating time series and cross section tourism demand models: Mainland United States to Hawaii data. *Tour. Manag.* **2011**, *32*, 28–38. [CrossRef]
21. Cheng, J.; Chen, Z. Impact of high-speed rail on the operational capacity of conventional rail in China. *Transp. Policy* **2021**, *110*, 354–367. [CrossRef]
22. Dobruszkes, F.; Chen, C.-L.; Moyano, A.; Pagliara, F.; Endemann, P. Is high-speed rail socially exclusive? An evidence-based worldwide analysis. *Travel Behav. Soc.* **2022**, *26*, 96–107. [CrossRef]
23. Ren, X.; Chen, Z.; Wang, F.; Dan, T.; Wang, W.; Guo, X.; Liu, C. Impact of high-speed rail on social equity in China: Evidence from a mode choice survey. *Transp. Res. Part A* **2020**, *138*, 422–441. [CrossRef]
24. Cavallaro, F.; Bruzzone, F.; Nocera, S. Spatial and social equity implications for High-Speed Railway lines in Northern Italy. *Transp. Res. Part A* **2020**, *135*, 327–340. [CrossRef]
25. Fang, J. Impacts of high-speed rail on urban smog pollution in China: A spatial difference-in-difference approach. *Sci. Total Environ.* **2021**, *777*, 146153. [CrossRef]
26. Jia, R.; Shao, S.; Yang, L. High-speed rail and CO_2 emissions in urban China: A spatial difference-in-differences approach. *Energy Econ.* **2021**, *99*, 105271. [CrossRef]
27. Strauss, J.; Li, H.; Cui, J. High-speed Rail's impact on airline demand and air carbon emissions in China. *Transp. Policy* **2021**, *109*, 85–97. [CrossRef]
28. Huang, Z.; Du, X. How does high-speed rail affect land value? Evidence from China. *Land Use Policy* **2021**, *101*, 105068. [CrossRef]
29. Okamoto, C.; Sato, Y. Impacts of high-speed rail construction on land prices in urban agglomerations: Evidence from Kyushu in Japan. *J. Asian Econ.* **2021**, *76*, 101364. [CrossRef]
30. Zhou, Z.; Zhang, A. High-speed rail and industrial developments: Evidence from house prices and city-level GDP in China. *Transp. Res. Part A* **2021**, *149*, 98–113. [CrossRef]
31. Tian, M.; Li, T.; Ye, X.; Zhao, H.; Meng, X. The impact of high-speed rail on service industry agglomeration in peripheral cities. *Transp. Res. Part D* **2021**, *93*, 102745. [CrossRef]
32. Xiao, F.; Lin, J.-J. High-speed rail and high-tech industry evolution: Empirical evidence from China. *Transp. Res. Interdiscip. Perspect.* **2021**, *10*, 100358. [CrossRef]
33. Chang, Z.; Diao, M.; Jing, K.; Li, W. High-speed rail and industrial movement: Evidence from China's Greater Bay Area. *Transp. Policy* **2021**, *112*, 22–31. [CrossRef]
34. Zhang, Y.; Ma, W.; Yang, H.; Wang, Q. Impact of high-speed rail on urban residents' consumption in China-from a spatial perspective. *Transp. Policy* **2021**, *106*, 1–10. [CrossRef]
35. Wang, X.; Huang, S.; Zou, T.; Yan, H. Effects of the high speed rail network on China's regional tourism development. *Tour. Manag. Perspect.* **2012**, *1*, 34–38. [CrossRef]
36. Jin, S.; Yang, J.; Wang, E.; Liu, J. The influence of high-speed rail on ice–snow tourism in northeastern China. *Tour. Manag.* **2020**, *78*, 104070. [CrossRef]
37. Wang, D.; Niu, Y.; Qian, J. Evolution and optimization of China's urban tourism spatial structure: A high speed rail perspective. *Tour. Manag.* **2018**, *64*, 218–232. [CrossRef]
38. Yin, P.; Lin, Z.; Prideaux, B. The impact of high-speed railway on tourism spatial structures between two adjoining metropolitan cities in China: Beijing and Tianjin. *J. Transp. Geogr.* **2019**, *80*, 102495. [CrossRef] [PubMed]
39. Zhang, A.; Liu, L.; Liu, G. High-speed rail, tourist mobility, and firm value. *Econ. Model.* **2020**, *90*, 108–116. [CrossRef]
40. Zhou, B.; Wen, Z.; Yang, Y. Agglomerating or dispersing? Spatial effects of high-speed trains on regional tourism economies. *Tour. Manag.* **2021**, *87*, 104392. [CrossRef]
41. Campa, J.L.; López-Lambas, M.E.; Guirao, B. High speed rail effects on tourism: Spanish empirical evidence derived from China's modelling experience. *J. Transp. Geogr.* **2016**, *57*, 44–54. [CrossRef]
42. Pagliara, F.; La Pietra, A.; Gomez, J.; Vassallo, J.M. High Speed Rail and the tourism market: Evidence from the Madrid case study. *Transp. Policy* **2015**, *37*, 187–194. [CrossRef]
43. Albalate, D.; Fageda, X. High speed rail and tourism: Empirical evidence from Spain. *Transp. Res. Part A* **2016**, *85*, 174–185. [CrossRef]
44. Albalate, D.; Campos, J.; Jiménez, J.L. Tourism and high speed rail in Spain: Does the AVE increase local visitors? *Ann. Tour. Res.* **2017**, *65*, 71–82. [CrossRef]
45. Guirao, B.; Campa, J.L. Cross effects between high speed rail lines and tourism: Looking for empirical evidence using the Spanish case study. *Transp. Res. Procedia* **2016**, *14*, 392–401. [CrossRef]
46. Guirao, B.; Campa, J.L.; Lopez-Lambas, M.F. The assessment of the HSR impacts on Spanish tourism: An approach based on multivariate panel data analysis. *Transp. Res. Procedia* **2016**, *18*, 197–204. [CrossRef]
47. Pagliara, F.; Mauriello, F.; Garofalo, A. Exploring the interdependences between High Speed Rail systems and tourism: Some evidence from Italy. *Transp. Res. Part A* **2017**, *106*, 300–308. [CrossRef]
48. Pagliara, F.; Mauriello, F. Modelling the impact of High Speed Rail on tourists with Geographically Weighted Poisson Regression. *Transp. Res. Part A* **2020**, *132*, 780–790. [CrossRef]

49. Masson, S.; Petiot, R. Can the high speed rail reinforce tourism attractiveness? The case of the highspeed rail between Perpignan (France) and Barcelona (Spain). *Technovation* **2009**, *29*, 611–617. [CrossRef]
50. Bugarin, M.R.; Novales, M.; Arcay, A. Alta velocitad y territorio. Algunas experiencias internacionales. *Ing. Y Territ.* **2005**, *70*, 4–11.
51. Pol, P. HST Station and Urban Dynamics: Experiences from Four European Cities. In *Railway Development*; Bruinsma, F., Pels, E., Rietveld, P., Priemus, H., van Wee, B., Eds.; Physica-Verlag HD: Heidelberg, Germany, 2008.
52. La Rocca, R.A. High Speed Train and Urban Tourism Attractiveness. *TeMA J. Land Use Mobil. Environ.* **2008**, *1*, 71–80.
53. Guerzoni, M. Alta Velocità e Piano per i Borghi. Così Rilanceremo il Turismo al Sud. Available online: https://www.corriere.it/politica/20_maggio_31/alta-velocita-piano-borghi-mosse-il-turismo-sud-d6945a08-a2ad-11ea-bc2b-bdd292787b00.shtml (accessed on 30 October 2021).
54. ISTAT. Movimento Turistico in Italia. 2019. Available online: https://www.istat.it/it/archivio/236148 (accessed on 30 October 2021).
55. Urban Index. 2013. Available online: https://www.urbanindex.it/indicatori/numero-di-luoghi-statali-della-cultura/ (accessed on 30 October 2021).
56. Mibact. 2017. Available online: http://www.statistica.beniculturali.it/ (accessed on 30 October 2021).
57. ISTAT. 2011. Available online: http://dati-censimentoindustriaeservizi.istat.it/ (accessed on 30 October 2021).
58. ISTAT. 2011. Available online: http://dati-censimentopopolazione.istat.it/Index.aspx (accessed on 30 October 2021).
59. Hansen, W.G. How accessibility shapes land use. *J. Am. Inst. Plan.* **1959**, *25*, 73–76. [CrossRef]