



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

## ARCHIVIO ISTITUZIONALE DELLA RICERCA

### Alma Mater Studiorum Università di Bologna Archivio istituzionale della ricerca

Innovation, productivity and spillover effects in the Italian accommodation industry

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

*Published Version:*

Innovation, productivity and spillover effects in the Italian accommodation industry / Bernini, Cristina; Galli, Federica. - In: ECONOMIC MODELLING. - ISSN 0264-9993. - ELETTRONICO. - 119:February(2023), pp. 106145.1-106145.14. [10.1016/j.econmod.2022.106145]

*Availability:*

This version is available at: <https://hdl.handle.net/11585/910022> since: 2022-12-20

*Published:*

DOI: <http://doi.org/10.1016/j.econmod.2022.106145>

*Terms of use:*

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

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

(Article begins on next page)

This is the final peer-reviewed accepted manuscript of:

**Cristina Bernini, Federica Galli. (2023). “Innovation, productivity and spillover effects in the Italian accommodation industry”. *Economic Modelling*, Vol. 119, February 2023, Art. No. 106145.**

The final published version is available online at:

<https://doi.org/10.1016/j.econmod.2022.106145>

Terms of use:

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

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

***When citing, please refer to the published version.***

# Innovation, Productivity and Spillover Effects in the Italian Accommodation Industry

Cristina Bernini

Department of Statistical Sciences, University of Bologna.

Via Belle Arti 41, Bologna 40126, Italy.

Ph. +390512098263

[cristina.bernini@unibo.it](mailto:cristina.bernini@unibo.it)

Federica Galli (*corresponding author*)

Department of Statistical Sciences, University of Bologna.

Via Belle Arti 41, Bologna 40126, Italy.

Ph. +390541434053

[federica.galli14@unibo.it](mailto:federica.galli14@unibo.it)

Declaration of Interest: none

**Abstract:** This paper aims at providing new micro-economic insights on how the different sources of innovation, both internal and external to the firm, contribute to determining the economic performance of Italian hotels, extending the current literature on industrial agglomeration to the accommodation sector at the firm level. To achieve this goal, we use georeferenced data on the consolidated accounts of Italian hotels for the period 2011-2019 and we estimate a spatial stochastic frontier model. Our results indicate that the performance of Italian hotels is mainly boosted by skilled labour and qualified human resources considering internal factors. On the other hand, we find that the innovative activity performed by neighbouring hotels spreads across space generating both agglomeration and competition effects. Our findings can be useful for policy makers and accommodation managers to improve hotels' production processes by taking advantage of innovative practices and spatial interactions.

**Keywords:** Innovation; Spillover effects; Spatial stochastic frontier models; Italian accommodation sector.

**Highlights**

- We provide new evidence on the link between innovation and hotels' performance
- We estimate a spatial stochastic frontier model using georeferenced firm level data
- Internally Italian hotels mainly rely on skilled labour and qualified human resources
- Hotels benefit from neighbours' investments in intangibles and human capital
- Registered patents and/or trademarks negatively affect neighbouring hotels

## **Acknowledgements**

We acknowledge the Editor, Prof. Sushanta Mallick, and three anonymous referees for their useful comments. They helped a lot to improve the quality of the paper.

## **1 Introduction**

Tourism is one of the most important engines for economic development (Brida et al., 2016; Pablo-Romero and Molina, 2013). Specifically, the positive link between tourism productivity and economic growth has been widely demonstrated by several authors (Schubert et al., 2011; Chang et al., 2012; Croes, 2013; Ridderstaat et al., 2016; Croes et al., 2021), leaving room for further research on how to achieve higher productivity levels in tourism in order to drive overall economic growth. As underlined by Poon (1993; 1994) and by Jackson and Murphy (2002; 2006), the key factor in the tourism sector is innovation, allowing both multinational companies and SMEs to remain competitive and achieve higher profitability levels. Indeed, in an increasingly competitive environment, it is becoming crucial for tourism firms<sup>1</sup> to innovate in order to attract new customer segments and add value to their tourism products and services (Marco-Lajara et al., 2016).

According to the Italian National Institute of Statistics (ISTAT) data from the "Main Annual Aggregates of National Accounts", tourism firms' investments in research and development and intellectual property products are still exiguous, being respectively the 0.36% and the 0.08% of the total investments performed by Italian firms in 2019. However, Italian tourism firms' investments in R&D and intellectual property products grew by 40.6% and 21.1% respectively between 2018 and 2019, 57.3% and 62.1% between 2017 and 2019, and 125.8% and 443.2% between 2011 and 2019. Therefore, despite the still insufficient amount of investments for innovation in the tourism sector, Italian tourism firms exceeded the national average

---

<sup>1</sup> Tourism firms include accommodation facilities, restaurants, travel and tourism agencies, attractions, transport companies, and handicraft shops.

growth rates for investments in R&D and patents in the last few years<sup>2</sup>, demonstrating that the actors of this sector are increasingly recognizing the importance of innovating to support the economic growth of the tourism industry.

Firms' innovative activity, competitiveness and success may be strongly related to investments in intangible assets (Montresor and Vezzani, 2016) because they allow new knowledge acquisition and process improvements. Intangible capital represents, between other things, the value of a company's information and communication technology (ICT), firm's organizational capital, and R&D investments. ICT application in accommodation facilities allows to speed up hotels' management procedures, upgrade the quality of economic operations, purchase tourist services online, communicate hotels' promotions and sales, recognize customers' profiles and offer personalized services, etc., and thus, it supports hotels' efficient functioning and competitiveness (Soteriades et al., 2004; Jaremen, 2016). On the other hand, R&D activity performed by hotels aims to raise the performance of existing operations by means of new or improved technologies, new job profiles, collaborative structures, and authority systems, to approach new markets and customer segments and to enable additional advantages to be offered to customers such as more comprehensive facilities and quality upgraded and speedier services (Hjalager, 2002). Examples of product and process innovation in the hotel industry concern environmentally sustainable practices, loyalty programmes, computerised management and monitoring systems, robots for cleaning and maintenance, self-service devices, electronic marketing, use of ICT in operations, automatic check-in and check-out, the introduction of touch-sensitive machines, virtual reality and smartphone apps, computerized reservation systems, technologies that ensure the mobility of people, luggage and goods such as x rays and iris-recognition, etc. (Hjalager, 2010; Jacob et al., 2010; Jacob and Groizard, 2007; Jimenez-Zarco et al., 2011). For a comprehensive review of hotels' innovative activity see Medina-Munoz et al. (2013).

Other than intangibles investments, in service sectors such as the accommodation industry, one of the most important sources of innovation is human resources (Chen et al., 2009). Good human resources management practices are positively associated with employee and customer satisfaction and with service quality, competitive advantage, better organizational performance, and lower turnover rates (Cho et al., 2006). Moreover, as demonstrated by Succurro and Boffa (2018), intellectual property is also a powerful and commonly used tool in

---

<sup>2</sup> Overall, Italian firms experienced an increase in investments in intellectual property products of 2.6% between 2018 and 2019, of 5.4% between 2017 and 2019, and of 28.9% between 2011 and 2019 while considering R&D, Italian firms' investments grew by 5.1%, 7.5%, and 21.2% in the same time periods.

the tourism sector aiming at developing a tourism brand strategy and at securing a competitive advantage. In this framework, more and more tourism firms rely on registered patents and/or trademarks to protect their innovative activity. Trademarks are useful for granting the owner the exclusive use of the brand while preventing its use by others, which protects valuable tools such as brands, logos, catchphrases, or slogans. In addition, patents are also useful to secure product innovations. Examples of hotels' patenting activity include door security systems, furniture modular systems, free-standing swimming pools, elevators with self-load bearing systems, bathroom aspirators, dehumidifiers for large rooms, and control systems for environmental pollution (Succurro and Boffa, 2018).

Despite the acknowledged importance of ICT, R&D, human resources and intellectual property strategies in increasing the performance of tourism firms, the tourism sector is characterized by low levels of research and development, lack of resources, rapid changes in ownership, high labour mobility, low salaries, low educational levels, and reluctance to take risks (Williams, Rodriguez Sanchez and Skokic, 2021). Moreover, the tourism sector is mainly composed of SMEs that hardly invest in R&D and therefore, due to the complexity of innovating inside the firm, the external environment in which hotels are embedded is a fundamental source of new knowledge. In this framework, scholars have largely acknowledged that the acquisition of knowledge from neighbours enhances innovativeness and competitiveness (Audretsch, 1998) both at the firm level (O'Mahony and Vecchi, 2009; Chyi, Lai and Liu, 2012) and at the local or regional level (Tappeiner et. al, 2008; Delgado, Porter and Stern, 2014; Koch and Simmler, 2020). Concentrating on the tourism sector, Sundbo, Orfila-Sintes, and Sørensen (2007) demonstrated that formal and informal relations among tourism firms, cooperation and networking contribute to determining the innovativeness of tourism firms. Similarly, analysing survey data from hospitality firms in Sweden, Backman, Klaesson, and Öner (2017) showed that cooperating with the other actors in the sector, such as suppliers, customers, competitors, and research organizations contributes to increasing innovation in the hotel sector. Moreover, Hameed, Nisar, and Wu (2021), investigating the link between external knowledge, internal innovation, and the performance of hotels situated in Pakistan, demonstrated that external knowledge and internal innovation positively affect firms' open innovation performance, leading to service innovation and increased business performance. In this framework, also Stojcic, Vojvodic, and Butigan (2019), analysing the performance of the Croatian hospitality industry during the period 2012-2014, found that knowledge and skills transferred through organizations foster

service innovation, confirming that the external sources of knowledge are fundamental drivers of innovation.

Thus, the external environment constitutes another fundamental source of new knowledge and innovation, but despite the recent and strong interest by researchers in investigating the impact of external sources of innovation on the productive outcome of hotels, a clear assessment of the magnitude, typology, and sources of these spatial effects is still lacking. Indeed, while scholars widely investigated the link between industrial agglomeration, innovation, and productivity from an empirical perspective considering the manufacturing, agricultural, and high-tech sectors, this topic is still relatively unexplored in the tourism sector (for a comprehensive review of previous studies on agglomeration and innovation see Carlin and Kerr (2015) and Binder (2019) for a focus on the tourism industry). However, for tourism-based countries such as Italy, being aware of the dynamics characterising neighbouring accommodation facilities is fundamental both for hotels and destination managers and for policy makers. Hence, this study aims at providing new micro-economic insights on how the different sources of innovation, both internal and external to the firm, contribute to determining the economic performance of Italian hotels, extending the current literature on industrial agglomeration to the accommodation sector at the firm level. In particular, we concentrate on the following research questions: (Q1) what sources of internal innovation contribute to increasing the productive performance of the Italian accommodation sector?; (Q2) what external innovative factors affect the productivity level of neighbouring tourism firms?; (Q3) does competition or agglomeration forces prevail?

To investigate the nature and the extent of agglomeration externalities occurring in the accommodation industry and affecting hotels' productivity and efficiency levels, we take advantage of a spatial stochastic frontier production function approach. The most appealing feature of stochastic frontier (SF) models is that they allow evaluating firms' performance by estimating a production frontier while simultaneously considering an inefficiency model distinguishing the random error from inefficiency. Recently, stochastic frontier models have been expanded in order to consider firms' clustering and interactions introducing some spatial components. Between existing spatial SF models, in this paper, we take advantage of the SDF-STE model developed by Galli (2021). The SDF-STE model allows to precisely disentangle three different kinds of spatial spillover effects occurring across neighbouring firms, i.e. productivity, inputs, and determinants of inefficiency spillovers, thanks to the introduction of the spatial lag of the dependent variable, of the input variables and the inefficiency determinants, respectively. Through this model, we are able to capture (i) the

overall level of global spatial dependence in the sector; (ii) spillover effects related to labour and capital; (iii) specific spatial effects related to the different sources of firms' innovative activity (human capital investments, intangible capital, patents and trademarks filing). The main characteristic of the SDF-STE consists in introducing the possibility of directly evaluating how each variable that determines the inefficiency level of neighbouring firms also affects nearby producers, giving rise to precise, detailed and distinct insights concerning spatial spillovers related to each source of internal innovation.

To perform our analysis, we take advantage of a georeferenced microdata sample belonging to the Italian ATECO55 sector collected from the AIDA Bureau Van Dijk database that contains budget data for the Italian accommodation firms over the period 2011-2019. In particular, we proxy innovation in the accommodation sector by using information on patents and trademarks, human capital exploitation, and intangible investments (see among others: Bernini and Guizzardi, 2010; Hameed, Nisar, and Wu, 2021). The use of georeferenced microdata allows for a precise assessment of micro-spatial patterns occurring at the firm level taking firms' heterogeneous characteristics into consideration.

To sum up, to the best of our knowledge, this is the first paper aiming at providing micro-economic evidence on the relationship between internal and external innovation and hotels' productivity using appropriate spatial econometric techniques. Specifically, estimating a novel spatial stochastic frontier model considering three different kinds of spatial effects, we are able to clearly assess the different spatial patterns occurring across nearby hotels that influence the productivity and efficiency level of neighbouring firms, where the former concerns how much output is generated from inputs and the latter refers to how well inputs are processed into output. Our findings show that both internal innovative activity related to intangible capital and investments in human resources significantly contribute to boosting hotels' performance. Moreover, externalities from neighbouring firms' innovative activity spread across space, generating positive feedback to nearby accommodation facilities. Therefore, although hotels' innovative activity is still limited and undervalued, hotels and destination managers should be aware of the influential role that different sources of internal and external innovation can play in strengthening the productivity of the entire sector. Policy makers can therefore rely on the empirical results from this paper to develop innovative and proper policies targeted at favouring the restart of the tourism sector after the stagnation period due to the current Covid-19 pandemic, exploiting spatial interactions characterizing Italian hotels.



## **2 Productivity, Knowledge Spillovers, and Agglomeration Externalities**

In the economic geography literature, three different kinds of spatial spillover effects occurring across clustered firms have long been recognized: productivity, inputs, and knowledge spillovers. Productivity spillovers can depend on emulation processes, meaning that less efficient producers try to emulate the best practices and procedures of the productivity leader in closely related industries to gain a productive advantage (Syverson, 2011). Considering input spillovers, the concentration of companies creates a shared market for workers with industry-specific skills and fosters the production of non-tradable specialized inputs, reducing costs and guaranteeing firms accessible and widely available specific products. Finally, the geographical concentration of firms stimulates innovative activity, spreading new knowledge through a tacit diffusion process. Several studies showed that knowledge spillovers play a relevant role in firms' industrial activity (Spence, 1984; Levin and Reiss, 1988; Adams and Jaffe, 1996; Griffith, Harrison, and van Reenen, 2006) and researchers generally agree on the idea that R&D investments can spread over a large number of productive units.

In recent tourism literature, tourism clusters have begun to be considered as a form of industrial clusters (Jackson and Murphy, 2002; Shaw and Williams, 2009). Therefore, since a tourist destination is composed of a conglomeration of competing and collaborating activities trying to cooperate to reach greater exposure and to build up a successful tourism product (Jackson and Murphy, 2006), tourism clusters, similar to manufacturing clusters, benefit from the existence of positive spillover effects resulting from spatial proximity, trust and shared values that encourage cooperation, social contact, and imitation (Shaw and Williams, 2009). As a consequence, clustered hotels experience higher productivity levels thanks to enhanced knowledge and innovation sharing (Adam and Mensah, 2013).

A few recent studies have analysed the connection between tourism clusters, internal and external knowledge, and hotel productivity, finding different and contrasting results. Peiró-Signes et al. (2015), analysing the impact of locating inside or outside U.S. tourism clusters on hotel economic performance using a concentration measure, suggested that belonging to a cluster improves a hotel's productivity. They demonstrated that the positive effect of agglomeration is more pronounced for luxury, upscale hotels and chain-managed hotels while it is less evident for resorts and airport locations. Differently, Baum and Mezias (1992), evaluating the impact of localized competition on hotels' failure rates for the Manhattan hotel industry in 1898-1990, found that hotels located in denser regions tend to experience higher failure rates. Taking internal and external knowledge into consideration, Marco-Lajara et al.

(2016), using multiple linear regression, showed that external knowledge coming from similar activities, universities and technological research institutions significantly contribute to hotel profitability. Marco-Lajara et al. (2019) demonstrated that agglomeration is positively associated with hotel profitability but with a lower effect than would be expected. Considering external knowledge, the authors did not find evidence of a significant positive effect of agglomeration on the acquisition of external knowledge, and consequently to profitability. Differently, using longitudinal data from lodging firms located in Southern China, Zhang et al. (2015) found that local entrepreneurs tend to imitate successful pioneering businesses. Hence, tacit knowledge spillovers among hotels located in the same region, generate a successful local development of the tourism sector considering an extended period. From a macroeconomic point of view, the tourism sector contributes considerably to the economic regeneration and development of an entire nation (Thomas and Long, 2001). Therefore, several studies concentrated on the effect of tourism agglomeration on local or national productivity. Investigating the impact of tourism agglomeration economies on UK regional productivity, Kim et al. (2021) found a positive effect of spatial agglomeration on hotel productivity due to knowledge spillovers and skilled labour pooling using a spatial panel data model. Moreover, Yang (2012), using a dynamic panel data model, showed a positive association between tourism agglomeration and the development of Chinese provinces in the period 2000-2009. In 2016, examining the impact of tourism agglomeration on labour productivity in Chinese provinces from 2000 to 2011, they also found a positive association between agglomeration density and productivity level but they showed that diversity of the tourism industry negatively affects labour productivity.

Nevertheless, to the best of our knowledge, no studies have yet investigated the link between tourism productivity, internal and external innovation from a micro-economic perspective using adequate spatial econometric techniques. The current studies have only focused on standard linear regression models, structural models, or concentration measures to investigate agglomeration externalities in tourism clusters leading to different and contrasting results. Our spatial stochastic frontier approach allows to separate the random error from inefficiency and to simultaneously estimate the frontier function and the efficiency model, distinguishing between productivity and efficiency determinants. Moreover, differently from other approaches, through the introduction of three different spatial terms, it is possible to capture global productivity spillovers as well as indirect effects related to the input variables and to the determinants of firms' inefficiency. Therefore, in this paper, we take advantage of a novel spatial stochastic frontier model to disentangle the different kinds of spatial spillover

effects affecting the economic performance of Italian hotels. The results from this study would be very useful for policy makers to design ad hoc place-based policies that exploit the existence and the magnitude of the different spatial effects characterizing nearby hotels to stimulate the productivity of the entire sector.

### 3 Methodology

#### 3.1 A New Proposal: the SDF-STE Model

Stochastic frontier (SF) models are a very commonly used tool in order to investigate firms' productive efficiency since they allow distinguishing the random disturbance from an inefficiency error component. However, basic SF models rely on the inappropriate assumption of spatial independence, since in fact, firms tend to cluster and share information. As a consequence, several authors started expanding the basic stochastic frontier model introducing some spatial components (Glass et. al, 2016; Tsukamoto, 2019; Orea and Alvarez, 2019; Galli, 2021). Between current spatial stochastic frontier models, in this analysis, we take advantage of the spatial Durbin stochastic frontier model for panel data introducing spillover effects in the determinants of firms' efficiency (SDF-STE) proposed by Galli (2021). This specification is the only one that allows (i) distinguishing between three different kinds of spatial spillover effects and (ii) evaluating the specific spatial effects arising from each source of neighbouring hotels' innovative activity thanks to the introduction of the spatial lag of each inefficiency determinant. The main advantage of this econometric approach concerns the possibility to compute direct and indirect effects related to the variables that identify hotels' innovative activity. This can be achieved by adding the exogenous spatial lag of the determinants of firms' efficiency in the efficiency model in the same fashion as a standard spatial lag model (SLX). The specification of the SDF-STE model is shown in Eq.(1)-(2) for  $i = 1, \dots, N$  and  $t = 1, \dots, T$ .

$$Y_{it} = X_{it}\beta + \rho \sum_{j=1}^N w_{ij} Y_{jt} + \sum_{j=1}^N w_{ij} X_{jt}\theta + v_{it} - u_{it} \quad (1)$$

$$\mu_{it} = Z_{it}\phi + \sum_{j=1}^N w_{ij} Z_{jt}\delta \quad (2)$$

Specifically,  $Y_{it}$  is the productive output of firm  $i$  at time  $t$  while  $X_{it}$  ( $1 \times k$ ) contains the  $k$  production inputs used by firm  $i$  at time  $t$  with associated parameter vector  $\beta$  ( $k \times 1$ ). Moreover, we consider spatial dependence in the frontier function introducing the spatial lag of the

dependent variable (SAR term) and the spatial lag of the input variables (SLX term). Specifically,  $\rho$  is the parameter that refers to the SAR term, capturing global spatial spillovers;  $\theta$  is the parameter vector ( $k \times 1$ ) associated with the SLX term capturing exogenous local spatial spillovers while  $w_{it}$  refers to the generic element in the  $i$ -th row and  $j$ -th column of the spatial weight matrix  $W$  containing non-negative spatial weights to identify neighbours (indexed by  $j = 1, \dots, N$ ) and elements equal to zero on the main diagonal. As is typical in stochastic frontier models, the error term is composed of two independent components:  $v_{it}$  and  $u_{it}$ . While  $v_{it}$  is the normally distributed random disturbance with zero mean and variance  $\sigma_v^2$ ,  $u_{it}$  represents technical inefficiency and is assumed to be distributed as a truncated normal random variable with a known mean  $\mu_{it}$  and variance  $\sigma_u^2$ . Following Battese and Coelli (1995), the two variance parameters are reparameterized as  $\sigma^2 = \sigma_u^2 + \sigma_v^2$  and  $\lambda = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$ . Moreover, the mean  $\mu_{it}$  of the technical inefficiency term  $u_{it}$  is modelled as a function of  $m$  exogenous variables  $Z_{it}$ , with associated parameter vector  $\phi$  ( $m \times 1$ ). The main characteristic of the SDF-STE model concerns the introduction of the spatial lag of the  $Z$  variables allowing to evaluate the specific spillover effects associated with the determinants of inefficiency of neighbouring hotels. Thus, the parameter vector  $\delta$  ( $m \times 1$ ) related to the spatial lag of the  $Z$  variables captures the different spillover effects originating from the inefficiency determinants of nearby producers and affecting firms' efficiency level.

The unknown parameters  $(\beta, \rho, \theta, \phi, \delta, \sigma^2, \lambda)$  of the SDF-STE model can be simultaneously estimated using a maximum likelihood estimation approach. In particular, while the spatial parameters related to the exogenous spatial lags of the  $X$  and of the  $Z$  variables can be straightforwardly estimated following standard ML estimation approaches, the introduction of the endogenous spatial lag of the dependent variable requires the introduction of the log determinant of the spatial filter  $(I_{NT} - \rho W)$  in the loglikelihood function. More details on the loglikelihood function and on the estimation procedure can be found in Appendix C. Testing various parameter restrictions through likelihood ratio tests or using the AIC or BIC information criteria, allows us to test whether it is better to consider this general model introducing different kinds of spatial effects or simpler nested specifications (Tsukamoto, 2019; Glass, Kenjegalieva, and Sickles, 2016; Adetutu et al., 2015; Battese and Coelli, 1995; Aigner, Lovell, and Schmidt, 1977).

### 3.2 The Empirical Model

We adopt a Translog specification to model the frontier production function defined in Eq.(1) due to the greater flexibility compared to a Cobb-Douglas specification. The specification of the empirical model is shown in Eq. (3) for  $i, j = 1, \dots, N$  ( $i \neq j$ ) and  $t = 1, \dots, T$ .

$$Y_{it} = \beta_0 + \beta_L L_{it} + \beta_K K_{it} + \beta_{LL} L_{it}^2 + \beta_{KK} K_{it}^2 + \beta_{LK} L_{it} K_{it} + \beta_t t + \beta_{2t} t^2 + \beta_{tL} t L_{it} + \beta_{tK} t K_{it} \quad (3)$$

$$+ \rho \sum_{j=1}^N w_{ij} Y_{jt} + \theta_L \sum_{j=1}^N w_{ij} L_{jt} + \theta_K \sum_{j=1}^N w_{ij} K_{jt} - u_{it} + v_{it}$$

Specifically,  $Y_{it}$  represents the logarithm of the value added of the hotel  $i$  at time  $t$  and the two production inputs,  $L_{it}$  and  $K_{it}$ , are, respectively, the logarithms of the number of employees and of fixed assets. Following Glass, Kenjegalieva, and Sickles (2016) we assume Hicks-neutral technical change and therefore the time trend variable  $t$  and its square are added to the model specification ( $t$  has a minimum value of 1 for the year 2011 and it increases by 1 for each year, reaching a maximum value of 9 for 2019). The interactions between time and the two input variables are also taken into consideration. The spatial lag of the dependent variable and the spatial lag of the two production inputs are introduced in the model to take global and local spatial dependence into account. Specifically,  $\rho$  captures global spatial dependence while  $\theta_L$  and  $\theta_K$  capture how the level of labour and capital of firm  $i$  is influenced by the input levels of neighbouring firms  $j$ , with  $j = 1, \dots, N$ . To identify neighbouring hotels, we use a time invariant row-standardized inverse distance spatial weight matrix, having all zeros on the main diagonal. Therefore,  $w_{ij}$  indicates the weight associated with each pair of spatial units  $i, j$  and it is equal to  $1/d_{ij}$  before normalization, where  $d_{ij}$  is the distance between the two units expressed in kilometres. While contiguity matrices are usually chosen when dealing with areal data, defining  $W$  as an inverse distance matrix is a common specification when working with points data as it allows to consider the relations of neighbours with all territorial units considering the exact spatial position of each element in the sample. Moreover, using a dense inverse distance matrix has several advantages. First of all, it implies not choosing an arbitrary truncation point or a cut-off for the number of neighbours so that subjective choices related to defining the neighbouring hotels do not affect the estimation results. Second, compared to inverse squared or polynomial distance matrices, a simple inverse distance matrix assumes that the relations between neighbouring observations are linear, which means that the strength of the relationship varies proportionally to the distance. Finally, with respect to matrices based on economic distance, it ensures that the spatial weights are exogenous.

However, in subsection 5.3 we test the robustness of our results with respect to alternative spatial weight matrices. In particular, we define various truncation points for  $W$  at a 200, 100, 50 and 30 kilometres radius around each spatial unit and we consider the 400, 250, 100, 50 and 30 nearest neighbours.

Finally,  $u_{it}$  is the inefficiency error term distributed as a truncated normal random variable with mean  $\mu_{it}$  and variance  $\sigma_u^2$  while  $v_{it}$  is the normally distributed error term with zero mean and variance  $\sigma_v^2$ , and  $u_{it}$  and  $v_{it}$  are assumed to be independent random variables. We model the mean  $\mu_{it}$  of firms' technical inefficiency as shown in Eq. (4).

$$\begin{aligned} \mu_{it} = & \phi_0 + \phi_{Hum}Hum_{it} + \phi_{Int}Int_{it} + \phi_{Pat}Pat_{it} + \phi_{Trad}Trad_{it} + \phi_{Size}Size_{it} + \phi_{DSize}DSize_{it} \quad (4) \\ & + \delta_{Hum} \sum_{j=1}^N w_{ij} Hum_{jt} + \delta_{Int} \sum_{j=1}^N w_{ij} Int_{jt} + \delta_{Pat} \sum_{j=1}^N w_{ij} Pat_{jt} + \delta_{Trad} \sum_{j=1}^N w_{ij} Trad_{jt} \\ & + \delta_{Size} \sum_{j=1}^N w_{ij} Size_{jt} + \delta_{DSize} \sum_{j=1}^N w_{ij} DSize_{jt} + \phi_{City}City_{it} + \phi_{Cult}Cult_{it} \\ & + \phi_{Sea}Sea_{it} + \phi_{Lake}Lake_{it} + \phi_{Mou}Mou_{it} + \phi_{CSea}CSea_{it} + \phi_{CMou}CMou_{it} \\ & + \phi_{More}More_{it} + \phi_{Nocat}Nocat_{it} + \phi_{Notur}Notur_{it} \end{aligned}$$

We assume that the mean of the technical inefficiency error term depends on determinants under the firms' control and on spillover effects coming from its neighbours. For the internal factors affecting hotel efficiency, we consider hotel size and firm innovative activity proxied by patents and trademarks filed, human capital exploitation, and intangible capital investments. To measure a hotel's innovative activity, we include the following variables in the model:  $Hum$ ,  $Int$ ,  $Pat$ , and  $Trad$  where  $Hum$  proxies firm investments in human capital, and is defined as the logarithm of the ratio between total annual labour costs and the number of employees. In the absence of data on the quality and education of workers for proxying human capital, firm income statement data can be considered the best approximation for measuring human resources value (Lev and Schwartz, 1971; Wyatt & Frick, 2010). Labour costs (including wages and training costs) per worker can be used as a proxy per human capital investments based on the assumption that firms with higher average labour costs per employee tend to recruit highly skilled workers (Martin and Moldoveanu, 2003; Wakelin, 1998; Le and Pomfret, 2011; Sari et al., 2016) since wages tend to vary more across firms for differences in human capital than because of worker rents (Pulakos et al., 2003; Hsieh and Klenow, 2011; Benkovskis, 2018). Besides salaries, incentives, study grants, awards and social security costs, labour cost measures generally include a substantial part of recruiting and

training costs because they are usually performed within the company by the firm's staff (Garcia-Ayuso et al., 2000). Moreover, Lajili and Zeghal (2006) demonstrated that, between other indices, indicators based on total labour expenditures are associated with higher abnormal returns, indicating that investors tend to perceive labour costs as a measure of human capital assets, rewarding it with greater market value. Thus, since the cost of investing in human capital is proportional to the cost of labour (Rhee and Pyo, 2010; Sydler et al., 2014), we can assume labour cost as a proxy variable of human capital.

Firms' investments in intangible capital (*Int*) are measured as the logarithm of the ratio between total capital (immaterial plus material) and fixed capital. Therefore, this variable equals 0 for hotels that do not make any investment in intangible capital while it shows increasing values as investments in immaterial capital increase. The rationale is that the propensity to invest in immaterial capital is strongly related to hotels' ability to introduce innovations because new investments are linked to the development of new technologies. Focusing on product innovation, patents are a very commonly used indicator because patenting allows innovative hotels to protect the newly developed product as trade secrets, giving the innovative firm a competitive advantage (Hameed, Nisar, and Wu, 2021). In this case, we introduce in the model a dummy variable (*Pat*) that equals 1 if, during the period considered (2011-2019), the hotel registered at least one patent, and 0 otherwise. Registered trademarks can be used as an additional indicator for service innovation (Gotsch and Hipp, 2012), since trademarks help protect highly valuable intangible assets, increasing a hotel's visibility and reputation (Marco-Lajara et al., 2016). In line with *Pat*, we measure registered trademarks by introducing a dummy variable (*Trad*) into the model that equals 1 if the hotel registered at least one trademark during the time period considered, and 0 otherwise. Finally, hotel size is proxied by the logarithm of the number of managers of the hotel (*Size*) in line with Bernini and Guizzardi (2010, 2016). In addition, enterprises driven by trained managers tend to use more capital and external finance and have different types of customers, encouraging innovative practices and new technology introduction (La Porta and Shleifer, 2008). Thus, besides measuring hotel size, the number of managers can also be considered an additional indicator of innovation other than intangibles investments, human capital, patents, and trademarks. Computing the logarithm implies obtaining missing values for those hotels with zero managers. Therefore, following the procedure suggested by Battese (1997), we substitute the missing values in *Size* with zero values and we take those hotels having zero managers into consideration including in the model a dummy variable (*DSize*) that equals 1 if the number of managers at the hotel is zero and 0 otherwise. However, we check the robustness

of our findings to a different definition of size using as alternative indicator the logarithm of total assets. The results shown in Table E5 of Appendix E confirm the robustness of our estimates.

In addition to hotel size and innovative activity, we also take hotel location into consideration due to the nature of the sector. Specifically, hotel location is taken into account through inclusion in the model of municipality dummy variables to identify the destination type according to the tourism municipality classification carried out by ISTAT in 2019. The dummy variable *City* is equal to 1 for hotels located in big cities with multidimensional tourism demand, *Cult* equals 1 for cultural, artistic, historical, or landscaped destinations, *Sea* for maritime destinations, *Lake* for lake destinations, *Mou* for mountain destinations, *CSea* for destinations that are both maritime and cultural, *CMou* for destinations that are both mountain and cultural, *More* for destinations that have more than two characteristics, *Nocat* for tourist destinations that cannot be categorized in this scheme, *Notur* for non-tourist destinations and *Therm* for thermal destinations (identified as the reference category). Besides the destination typology, we also considered including other location specific variables at the municipal level in the inefficiency model to capture specific territorial features. However, as shown in Table E4 of Appendix E, none of them was found to be significant at a significance level of 1%, so we excluded them from the final model specification.

Focusing on external influences, we also consider if and how the factors that contribute to determining neighbouring hotels' efficiency level also affect the level of efficiency of a given hotel. Introducing the spatial lag of multiple indicators such as *Size*, *DSize*, *Hum*, *Int*, *Pat*, and *Trad* into the model, we strengthen our analysis by considering different sources of spatial diffusion as suggested by Nelson (2009). In particular, the unknown parameters  $\delta_{Size}$ ,  $\delta_{DSize}$ ,  $\delta_{Hum}$ ,  $\delta_{Int}$ ,  $\delta_{Pat}$ , and  $\delta_{Trad}$  capture spillover effects resulting from being located near a big facility, near a hotel making large investments in human capital and/or in intangibles or near highly innovative hotels that have registered patents or trademarks. Therefore, modelling the mean of technical inefficiency as in Eq.(4) we are able to detect both direct and indirect effects affecting the efficiency level of Italian hotels.

### 3.3 Marginal Effects

It is well-known that the parameter estimates obtained from the spatial models cannot be interpreted in a meaningful way because they do not represent marginal effects. Accordingly, when the spatial lag of the dependent variable is included in the model, this endogenous interaction enters the first derivatives computation and the  $\beta$  estimates no longer represent



marginal effects. The first derivatives of the dependent variable with respect to labour ( $L$ ) and capital ( $K$ ), referring to a Translog production function, are shown in Eq.(5)-(6) respectively, using matrix notation. Specifically,  $I_{NT}$  refers to an  $(NT \times NT)$  identity matrix while  $1_{NT}$  is an  $(NT \times 1)$  vector of ones.

$$\frac{dY}{dL} = (I_{NT} - \rho W)^{-1} \left( I_{NT} \cdot \left( 1_{NT}^T \otimes (\beta_L 1_{NT} + 2\beta_{LL}L + \beta_{LK}K + \beta_{tLt}) \right) + W\theta_L \right) \quad (5)$$

$$\frac{dY}{dK} = (I_{NT} - \rho W)^{-1} \left( I_{NT} \cdot \left( 1_{NT}^T \otimes (\beta_K 1_{NT} + 2\beta_{KK}K + \beta_{LK}L + \beta_{tKt}) \right) + W\theta_K \right) \quad (6)$$

Starting from the two matrices obtained from the right-hand side of Eq. (5)-(6), direct, indirect, and total effects can be calculated following the method proposed by LeSage and Pace (2009). Direct effects can be found as the average of the diagonal element of the matrix on the right hand side of Eq.(5) and Eq.(6), indirect effects can be defined as the average of the sum of non-diagonal elements of those matrices, while total effects correspond to the sum of the previous two.

Similar to the  $\beta$  estimates, the  $\phi$  estimates of the inefficiency model cannot be interpreted as elasticities due to the influence of the spatial lag of  $Y$ . The first derivative of  $u$  with respect to a generic determinant  $Z$ , is shown in Eq. (7). Starting from the matrix obtained from the right-hand side of Eq. (7), the marginal effects associated with a generic determinant  $Z$  can be straightforwardly calculated, following the same procedure as described above.

$$\frac{du}{dZ} = (I_{NT} - \rho W)^{-1} (I_{NT} \phi_Z + W \delta_Z) \quad Z = Hum, Int, Pat, Trad, Size \quad (7)$$

#### 4 Data, Variables and Descriptive Statistics

The data used for the analysis were collected from the AIDA-Bureau Van Dijk database, which provides information on the consolidated accounts of Italian companies. We concentrated on the ATECO 55 sector, which refers to the Italian accommodation industry, in the time period 2011-2019 (ATECO is the Italian version of the NACE classification of economic activities). Starting from a sample of more than 20,000 individual observations available yearly, we ended up with a balanced panel of 5409 firms for each year due to the necessary cleaning procedure (Appendix A contains more details on the data cleaning procedure). Comparing our sample with population data retrieved from the Industry and Services Census carried out

by ISTAT in 2011, our sample is determined to be a good representation, covering 12.69% of total tourism firms and 35.72% of the total number of employees working in this sector. More details on the sample coverage rates by class of employees and macroarea, and for aggregated data at the municipal level are shown in Appendix B.

Table 1 describes all the variables used in the analysis (i.e. output, inputs, and inefficiency determinants) and shows some descriptive statistics. Only 2.1% of hotels from our sample are very small hotels (i.e. hotels with number of managers equal to zero) and this is due to the data cleaning procedure that excluded the very small enterprises from our sample (i.e. hotels with value added, fixed capital and personnel costs less than one thousand euros and number of employees less than one). Concentrating on innovative activity, it can be observed that the distribution of *Hum*, proxying human capital, is very concentrated between the 10th and the 90th percentiles, with a mean cost for employee of about 22 thousand euros per year. This variable reflects hotel managers' propensity to invest few resources in human capital, pay low wages, invest little money in incentives, awards and study grants, and hire employees with low educational levels. Moreover, many hotels from our sample invest very little money in intangible capital, indeed, *Int* has zero value for 19.65% of units in our sample while the median annual expenditure in intangibles equals 18 thousand euros overall and 37 thousand euros for hotels reporting positive investments in intangibles. Nevertheless, the distribution of *Int* is positively skewed and the 90th percentile equals 425 thousand euros per year, indicating that hotels' innovative activity is higher than usually believed. In addition, 16.8% of the hotels in our sample have registered at least one patent and 12.1% at least one trademark, in the time period considered. These statistics remark the low levels of research and development, salaries, and educational attainment in this sector due to a lack of resources, rapid changes in ownership, high labour mobility, and reluctance to take risks (Weidenfeld, Williams, and Bultler, 2010).

**Table 1:** Variables Description

Variable	Definition	Min	10th Perc.	Mean	90th Perc.	Max	SD
Y	$\log(\text{Value Added})$	0.001	4.23	5.78	7.32	12.01	1.25
L	$\log(\text{Number Employees})$	0	0.69	2.16	3.50	7.47	1.08
K	$\log(\text{Fixed Capital})$	0	3.47	6.37	8.86	13.40	2.18
t	Time	1	1	5	9	9	2.58
Hum	$\log\left(\frac{\text{Personnel Costs}}{\text{Number of Employees}}\right)$	0	2.25	3.10	3.76	7.56	0.73

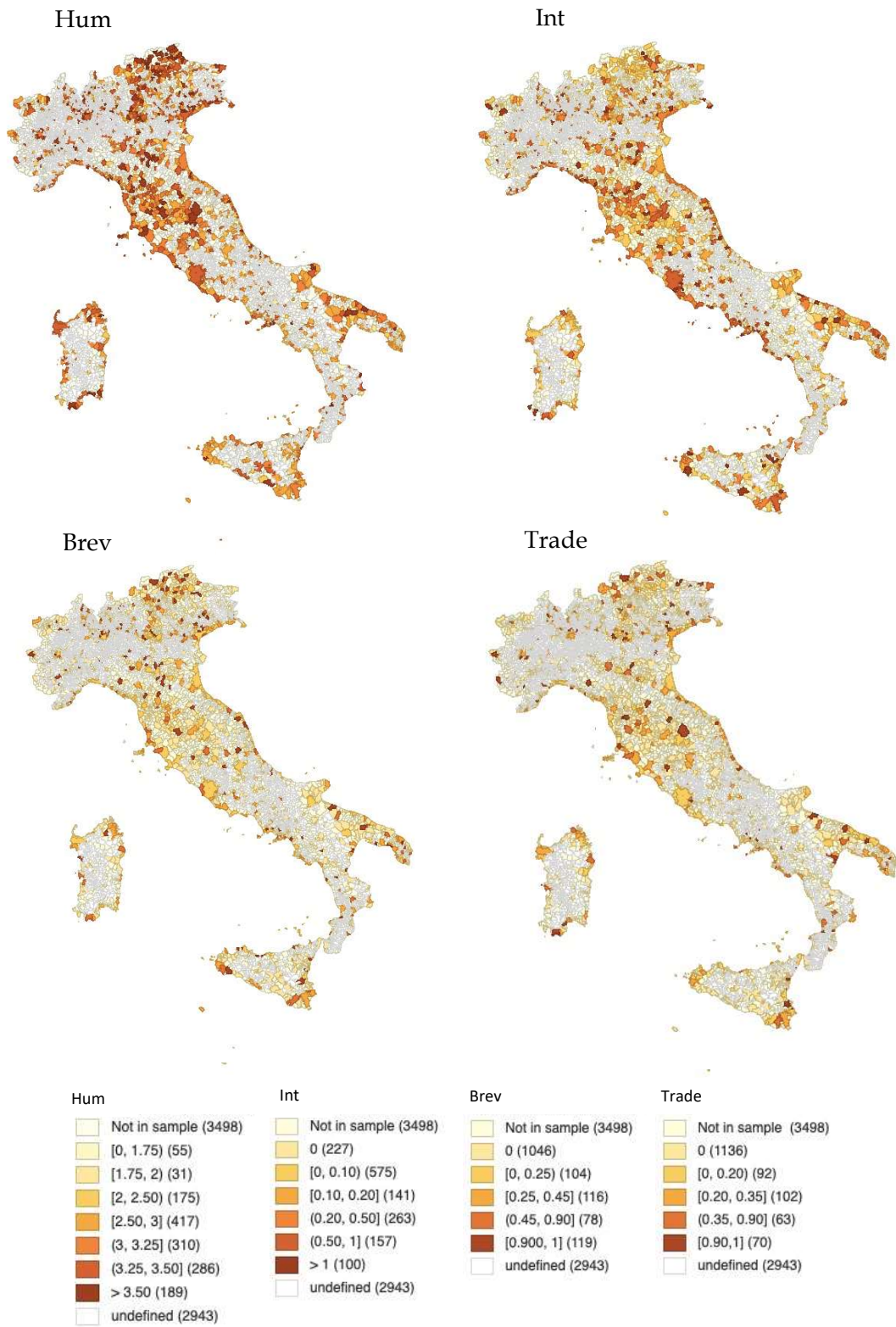
Int	$\log\left(\frac{\text{Total Capital}}{\text{Fixed Capital}}\right)$	0	0	0.31	3.76	7.67	0.60
Pat	1 if PatentRights > 0	0	-	0.17	-	1	0.37
Trad	1 if RegisteredTrademarks > 0	0	-	0.12	-	1	0.33
Size	$\log(\text{NumberManagers})$	0	0	0.54	1.39	3.76	0.67
DSize	1 if NumberManagers=0	0	-	0.02	-	1	0.14
City	1 if BigCity	0	-	0.20	-	1	0.40
Cult	1 if Cultural	0	-	0.10	-	1	0.29
Sea	1 if Sea	0	-	0.16	-	1	0.36
Lake	1 if Lake	0	-	0.03	-	1	0.17
Mou	1 if Mountain	0	-	0.03	-	1	0.16
CSea	1 if Cultural&Sea	0	-	0.20	-	1	0.40
CMou	1 if Cultural&Mountain	0	-	0.06	-	1	0.23
Therm	1 if Thermal	0	-	0.05	-	1	0.12
More	1 if MoreThanTwoVocations	0	-	0.06	-	1	0.24
Nocat	1 if NotCategorizable	0	-	0.12	-	1	0.32
Notur	1 if NonTouristDestination	0	-	0.01	-	1	0.07

Figure 1 shows how firms' innovative activity is distributed across the Italian territory at the municipal level. The "undefined" category refers to municipalities that do not contain any hotels while the "not in sample" category refers to municipalities that are not covered by our sample. Considering the remaining municipalities, Figure 1 shows that hotels investing in human capital are predominantly located in the North of Italy (specifically in Trentino Alto Adige, on the coast of Veneto, and in Emilia-Romagna) and in the Centre of Italy, mainly in Tuscany, Umbria, and Lazio. Concentrating on the South of Italy, the Apulia region, the South of Sicily, and the Northern and Southern coast of Sardinia are the areas where hotels make larger investments in human capital. The distribution of *Int* is similar to the one of *Hum* across Italy. The main difference is found in the area of Trentino Alto Adige and generally in all of North-East of Italy where hotels do not make as many investments in intangible capital as they do in human resources. Conversely, hotels located on the Campania coast invest more in intangibles than in human capital. Finally, in most municipalities, hotels did not register any patents (1046 municipalities) or trademarks (1136 municipalities). As for *Hum* and *Int*, patenting is more common in some municipalities in the North-East of Italy, in the Apulia region, and on the West coast of Sicily while municipalities where hotels register trademarks more frequently are mostly located in Trentino Alto Adige, in the Centre of Italy, in Apulia, and on the Southern tip of Sicily and Sardinia.

Figure 2 shows the Local Moran significance cluster map (LISA) at a significance level of 5%, which allows local geographical clusters to be identified and the degree of local spatial

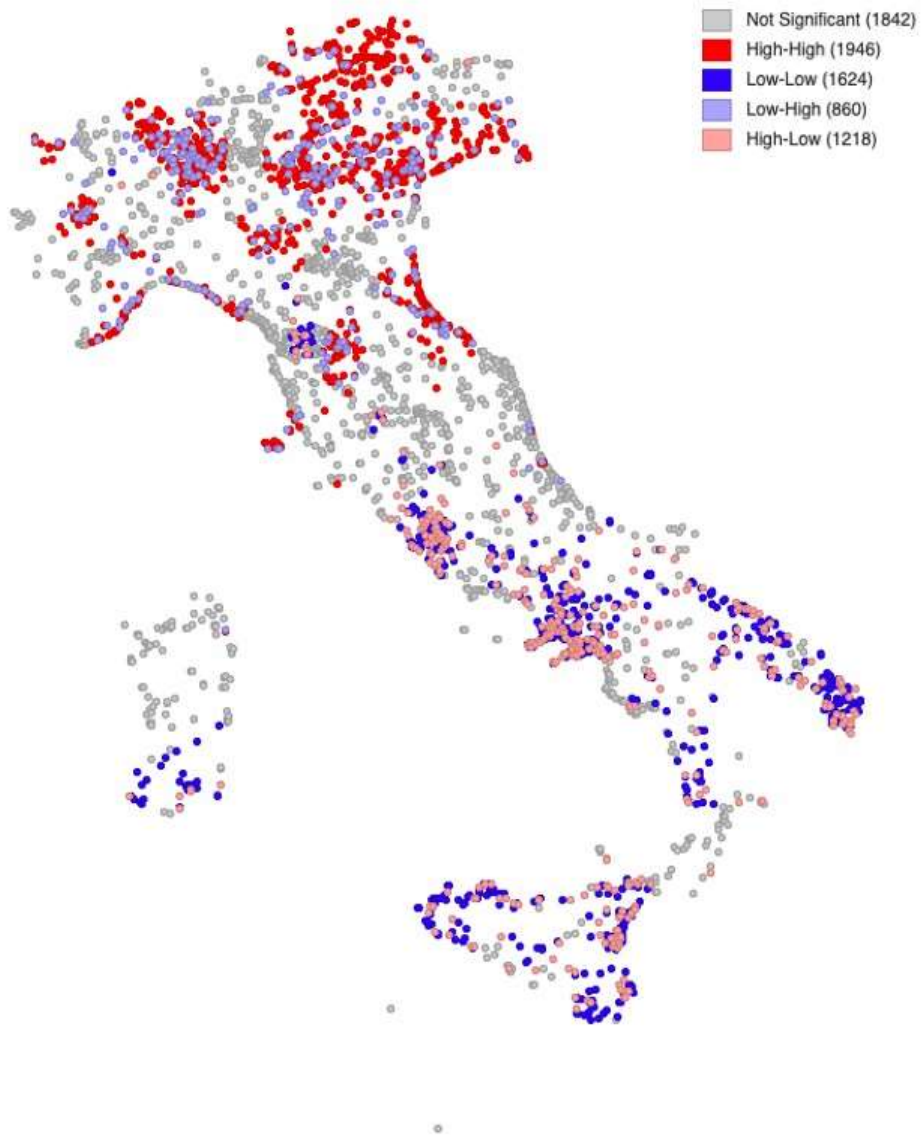
dependence to be determined. Considering all the observations having a value added greater than zero in the year 2019 (7508), only 1842 hotels are not affected in any way by local spatial dependence at a significance level of 5%. On the other hand, local clusters located in the North-Centre of Italy are high-high and low-high clusters, indicating that hotels in this area are mainly surrounded by hotels having a high value added. Conversely, the South-Centre of Italy is characterized by the presence of low-low and high-low clusters, suggesting that hotels located in the South-Centre of Italy are mostly near other hotels with low levels of value added. The area around Florence, which has significant low-low and high-low clusters, is the only exception to this clear separation into two areas characterized by two distinct typologies of local clusters (i.e. high-high and low-high clusters in the North-Centre and low-low and high-low clusters in the South-Centre of Italy).

**Figure 1.** Innovation across Italian Municipalities





**Figure 2.** LISA Significance Cluster Map: Value Added 2019  
(5% Significance Level)



## 5. Results

### 5.1 Estimation Results and Model Selection

Table 2 shows the estimation results of the SDF-STE model and of all the nested models, starting from the two non-spatial specifications (SF and SF-TE), passing to spatial models that do not take the determinants of technical inefficiency into consideration (SLXF, SARF, and SDF) and ending with the SARF-TE that includes a model for the determinants of firm inefficiency but that only considers the spatial lag of  $Y$  as spatial effect. Comparing the estimation results of the nested models, the estimated  $\beta$  and  $\phi$  coefficients are quite robust to different model specifications. Nevertheless, these coefficients cannot be interpreted in a meaningful way when the spatial lag of  $Y$  is included because they no longer represent simple partial derivatives. Thus, the related marginal effects are discussed separately in paragraph 5.2. Finally, considering the functional form of the production frontier, the result of the LR test indicates rejecting the null hypothesis of reducing the model to a simpler Cobb Douglas specification at a 1% significance level (the test statistic equals 2214.9). Thus, our Translog specification does not reduce to a Cobb Douglas one.

Concentrating on the spatial autoregressive parameter, the estimates of  $\rho$  are positive and significant at the 1% significance level across all the models introducing the SAR term, indicating that positive spillover effects occur at the global level in the Italian accommodation sector. Moreover,  $\rho$  appears to be almost doubled if the determinants of firm efficiency are not included in the model specification, in fact, it equals 0.352 using the SARF and 0.168 with the SARF-TE model while it equals 0.604 and 0.351 using the SDF and the SDF-STE specification, respectively. Indeed, as observed by Tsukamoto (2019) for the SARF and the SARF-TE models, when the determinants of firm efficiency are not considered in spatial stochastic frontier models, the spatial autoregressive parameter absorbs some of the heterogeneity related to technical inefficiency and it tends to be overestimated. To choose between the different nested models, different criteria can be used, such as the Akaike Information Criteria (AIC), the Schwarz/Bayesian Information Criteria (BIC) or alternatively, some Likelihood Ratio Tests for nested models can be implemented. Looking at the values of the AIC and of the BIC information criteria contained in Table 3 referring to all the estimated nested models, the model specification that minimizes both criteria is the SDF-STE. Additionally, the Likelihood Ratio Test always rejects the null hypothesis of reducing the number of parameters of the SDF-STE model in favour of a simpler specification. Therefore, the SDF-STE is the preferred model.



**Table 2: Nested Models Results**

	SF		SF-TE		SLXF		SARF		SDF		SARF-TE		SDF-STE	
	Coef.	t-stats	Coef.	t-stats	Coef.	t-stats	Coef.	t-stats	Coef.	t-stats	Coef.	t-stats	Coef.	t-stats
$\beta_0$	3.635***	38.63	6.665***	7.68	2.929***	30.17	1.630***	6.46	1.404***	16.06	6.095***	15.41	5.467***	13.71
$\beta_L$	0.633***	55.06	0.665***	85.11	0.629***	55.20	0.626***	61.35	0.628***	60.42	0.660***	83.53	0.662***	84.87
$\beta_K$	0.070***	10.64	0.077***	17.04	0.071***	10.95	0.073***	12.03	0.074***	12.11	0.077***	17.13	0.077***	17.20
$\beta_{LL}$	0.090***	42.76	0.057***	40.43	0.089***	44.50	0.089***	49.33	0.090***	49.78	0.057***	40.50	0.057***	40.57
$\beta_{KK}$	0.019***	31.50	0.010***	25.00	0.019***	31.17	0.019***	37.00	0.019***	37.20	0.010***	25.00	0.010***	22.50
$\beta_{LK}$	-0.050***	-33.40	-0.029***	-26.46	-0.050***	-33.60	-0.051***	-36.43	-0.051***	-36.57	-0.030***	-26.82	-0.029***	-29.00
$\beta_t$	0.067***	10.08	0.005	0.96	0.082***	12.36	0.059***	9.22	0.024***	3.97	0.002	0.43	-0.005	-1.09
$\beta_{2t}$	-0.007***	-14.00	-0.001**	-2.75	-0.009***	-17.40	-0.007***	-11.83	-0.004***	-7.60	-0.001***	-3.00	-0.001*	-1.25
$\beta_{tL}$	0.006***	5.08	0.004***	4.88	0.006***	5.25	0.006***	6.40	0.006***	5.64	0.004***	5.00	0.004***	5.00
$\beta_{tK}$	0.001**	2.00	0.001***	3.50	0.001**	1.83	0.001**	2.00	0.001**	1.80	0.001***	3.25	0.001***	3.00
$\rho$	-	-	-	-	-	-	0.352***	10.14	0.604***	31.30	0.168***	8.48	0.351***	18.30
$\theta_L$	-	-	-	-	0.175***	6.98	-	-	-0.340***	-15.10	-	-	-0.216***	-11.33
$\theta_K$	-	-	-	-	0.049***	4.64	-	-	-0.069***	-7.20	-	-	0.006	0.78
$\phi_0$	-	-	5.363***	6.22	-	-	-	-	-	-	5.748***	14.64	5.251***	13.23
$\phi_{hum}$	-	-	-0.653***	-210.74	-	-	-	-	-	-	-0.646***	-208.42	-0.647***	-208.58
$\phi_{Int}$	-	-	-0.116***	-29.05	-	-	-	-	-	-	-0.116***	-29.82	-0.115***	-29.54
$\phi_{pat}$	-	-	-0.054***	-9.38	-	-	-	-	-	-	-0.054***	-9.31	-0.054***	-9.53
$\phi_{trad}$	-	-	-0.039***	-5.91	-	-	-	-	-	-	-0.037***	-5.55	-0.038***	-5.78
$\phi_{size}$	-	-	-0.040***	-11.54	-	-	-	-	-	-	-0.037***	-10.46	-0.037***	-10.57
$\phi_{dsize}$	-	-	0.069***	4.73	-	-	-	-	-	-	0.077***	5.33	0.081***	5.66
$\delta_{hum}$	-	-	-	-	-	-	-	-	-	-	-	-	0.160***	7.31
$\delta_{Int}$	-	-	-	-	-	-	-	-	-	-	-	-	-0.072***	-2.71
$\delta_{pat}$	-	-	-	-	-	-	-	-	-	-	-	-	0.099***	2.80
$\delta_{trad}$	-	-	-	-	-	-	-	-	-	-	-	-	0.039	0.97
$\delta_{size}$	-	-	-	-	-	-	-	-	-	-	-	-	-0.013	-0.70
$\delta_{dsize}$	-	-	-	-	-	-	-	-	-	-	-	-	-0.358***	-3.85
$\phi_{city}$	-	-	-0.070***	-6.50	-	-	-	-	-	-	-0.076***	-7.04	-0.064***	-5.94
$\phi_{cult}$	-	-	0.041***	3.56	-	-	-	-	-	-	0.008	0.73	0.016*	1.43
$\phi_{sea}$	-	-	-0.071***	-6.57	-	-	-	-	-	-	-0.109***	-10.07	-0.095***	-8.89
$\phi_{lake}$	-	-	-0.149***	-9.64	-	-	-	-	-	-	-0.160***	-10.36	-0.139***	-9.06
$\phi_{mou}$	-	-	0.015	0.96	-	-	-	-	-	-	-0.017	-1.09	-0.007	-0.45
$\phi_{csea}$	-	-	-0.057***	-5.39	-	-	-	-	-	-	-0.099***	-9.47	-0.085***	-8.12
$\phi_{cmou}$	-	-	-0.058***	-4.50	-	-	-	-	-	-	-0.075***	-5.89	-0.065***	-5.04
$\phi_{more}$	-	-	-0.040***	-3.17	-	-	-	-	-	-	-0.064***	-5.15	-0.047***	-3.77
$\phi_{notcat}$	-	-	0.072***	6.52	-	-	-	-	-	-	0.041***	3.69	0.052***	4.71
$\phi_{notur}$	-	-	0.051*	1.60	-	-	-	-	-	-	0.002	0.06	0.015	0.48
$\phi_{therm}$	-	-	omit.	-	-	-	-	-	-	-	omit.	-	omit.	-
$\sigma^2$	0.705	-	0.203	-	0.694	-	0.678	-	0.674	-	0.200	-	0.199	-
$\lambda$	0.621	-	0.863	-	0.614	-	0.613	-	0.621	-	0.882	-	0.879	-
TE	0.64	-	0.67	-	0.64	-	0.64	-	0.64	-	0.61	-	0.62	-

\*\*\* :  $pvalue \leq 0.01$ ; \*\* :  $pvalue \leq 0.05$ ; \* :  $pvalue \leq 0.10$ ; omit=omitted

**Table 3:** AIC, BIC and Likelihood Ratio Tests

Model	LL	AIC	BIC	LL Test: H0	Constraints	Test Stat.	Decision
SF	-48071.5	96167	96272	$\delta, \phi, \theta, \rho = 0$	26	36419.2	Reject
SF-TE	-30197.2	60452	60707	$\delta, \theta, \rho = 0$	8	670.5	Reject
SLXF	-47905.0	95838	95961	$\delta, \phi, \rho = 0$	24	36086.3	Reject
SARF	-47414.2	94854	94968	$\delta, \phi, \theta = 0$	25	35104.6	Reject
SDF	-47221.1	94472	94604	$\delta, \phi = 0$	23	34718.4	Reject
SARF-TE	-29944.6	59949	60213	$\delta, \theta = 0$	8	165.4	Reject
SDF-STE	-29861.9	59799	60134	-	-	-	-

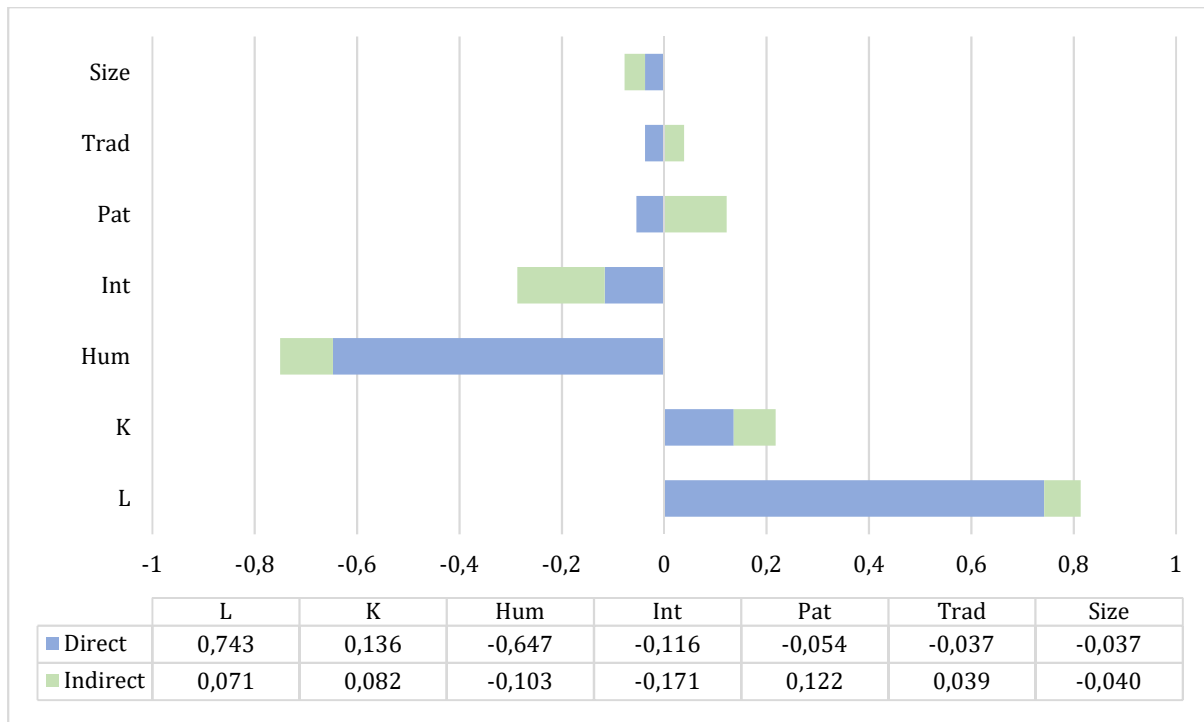
### 5.2 Marginal Effects

Focusing on the marginal effects related to the preferred SDF-STE model, Figure 3 shows the direct and indirect effects associated with the two input variables ( $L, K$ ) and with the determinants of firm efficiency ( $Hum, Int, Pat, Trade, Size$ )<sup>3</sup>. Further insights on the related p-values and on the total effects can be found in the first column of Table 4.

**Figure 3.** Marginal Effects

---

<sup>3</sup> For the inefficiency determinants, negative effects should be interpreted as decrease in inefficiency levels while positive impacts relate to an increase in hotels' inefficiency.



The direct effects of labour and capital on hotels' value added are equal to 0.743 and 0.136 respectively, while the indirect effects equal 0.071 and 0.082, respectively. Therefore, while the direct effect of labour is greater in magnitude than the one of capital (i.e. the accommodation sector is a labour-intensive industry), the indirect effect of capital is slightly higher than the one of labour, meaning that having bigger hotels as neighbours positively influences firm productivity level more than having neighbouring hotels that invest in labour. The total effects of labour and capital are equal to 0.814 and 0.218, respectively. Hence, the return to scale parameter equals 1.032 (significantly different from 1 at a 5% significance level), indicating the presence of increasing returns to scale. Considering technical change over years, the coefficients related to time shown in the first column of Table 4 are positive and significant considering both direct, indirect, and total effects, indicating that the production frontier shifts upward over time thanks to technological change.

Moving to the determinants of hotels' efficiency, *Hum*, *Int*, *Pat*, *Trad*, and *Size*, all have a negative and significant direct effect on firm inefficiency level indicating that all the different sources of internal innovation considered in this study, as well as size, positively influence hotels' performance (Q1). Human capital (-0.647) is the factor that contributes most to positively affecting the efficiency level of Italian hotels followed by investments in intangible capital (-0.116). Indeed, human capital is a key source of innovation in the accommodation industry because of the intangible nature of this sector and the simultaneity of production and consumption in service delivery (Ottenbacher, 2007). Specifically, in order to obtain a 10%

increase in efficiency, Italian hotels need to invest about 215 thousand euros in intangibles or alternatively, to increase labour investments per worker by around 3.60 thousand euros yearly. Moreover, in line with the results of Orfila-Sintes et al. (2005), we find that hotels' size is positively associated with efficiency as larger firms have the advantage of economies of scale in innovation activities (Camisón-Zornosa et al., 2004). Finally, also the direct effects related to registered patents and trademarks result to be positively associated with hotels' efficiency level, increasing firm visibility and protecting product and service innovations (Marco-Lajara et al., 2016).

Considering the indirect effects, our findings indicate that all the different sources of innovation considered in this study significantly impact neighbouring accommodation facilities (Q2) but with different effects (Q3). Indeed, while we detect positive spillovers related to intangibles, size, and human capital investments, innovative activity associated with patents and trademarks generates negative spatial effects among neighbouring hotels. Among the sources of innovation that generate positive feedback, *Int* is the variable that contributes the most to positively affecting neighbours (-0.171). The magnitude of the coefficient related to the indirect marginal effect of intangibles is greater than the one associated with the direct effect, indicating that overall, positive spillovers generating from innovative activity overcome direct internal effects. According to our results, hotels benefit more from investments in intangible capital performed by neighbours than from internal investments highlighting the fundamental role played by the few innovators in the sector as knowledge and innovation disseminators. Thus, despite the difficulty of innovating inside the firm due to the peculiarities of this industry, in the accommodation sector, it is fairly easy to adopt new knowledge coming from neighbours because the operational processes are quite evident and also the technological level is basic (Weidenfeld et al., 2010, Decelle, 2006; Hjalager, 2002). Besides having neighbours performing an intense innovative activity, also having nearby hotels that invest in human capital generates a positive and significant spillover effect (-0.103) but differently from *Int*, the direct effect associated with *Hum* greatly exceeds the indirect one. Therefore, positive feedback effects also generate from skilled human resources in neighbouring accommodation facilities thanks to social contact, shared ideas between individuals, learning by observation, human relationships, and imitation (Yang, 2012). However, internal investments in human capital retain a key role in this sector due to their direct connection with customer satisfaction, service quality, and better organizational performance which in turn leads to increased hotel performance (Cho et al., 2006). Finally, considering hotel size, the estimated coefficient for the indirect marginal effect of *Size* indicates

that, in line with the positive indirect effect related to capital, having big hotels as neighbours positively influences a hotel's performance. Indeed, bigger hotels tend to be more innovative with respect to smaller and medium-sized hotels generating positive spillover effects that are beneficial to all neighbouring accommodation facilities.

On the other hand, the indirect effects of *Pat* and *Trad* on hotels' inefficiency are both significant and positive (0.122 and 0.039, respectively), indicating that hotels are disadvantaged when neighbouring firms registered trademarks or patents in the previous years. Therefore, the protective function performed by patents and trademarks is found to be effective, because, in addition to providing innovative firms a productivity advantage, it also weakens neighbouring hotels through negative spillover effects. These results are in line with Haschka and Herwartz (2020) who demonstrated that patent blocking might be crucial for innovative firms to strategically secure their technological expertise, generating negative competitive spillovers. Specifically, they showed that there is a negative association between the successful performance of competitors and the efficiency of the innovative process of peers, proxied by patents. Indeed, patents give innovative firms the exclusive right to commercialize the newly patented products for a certain period of time. Similarly, registered trademarks are used to protect hotels' highly valuable intangible assets and to differentiate firm services from potentially competing services (Hameed, Nisar, and Wu, 2021).

To sum up, our results suggest that while the performance of Italian hotels is mainly boosted by skilled labour and qualified human resources considering internal factors (Q1), positive spatial effects are primarily linked to capital, and in particular, to intangible capital, as far as external factors are concerned (Q2). Finally, while positive feedbacks arise from intangibles, hotels' size and human capital investments, patent and trademark activity performed by neighbours generate competition effects (Q3).

### ***5.3 Does distance matter in shaping agglomeration externalities?***

The effect of agglomeration externalities can vary depending on the spatial distance considered (Arbia, 1989). Hence, it is interesting to evaluate if the magnitude of the spatial effects detected earlier is robust to different specifications of the spatial weight matrix. This further analysis allows us to precisely identify how indirect effects are affected by the geographical distance considered. This information can be very relevant to policy makers in understanding how to combine general and regional-specific policies exploiting the existence of spatial interactions in line with Cabres-Borras and Serrano-Domingo (2007).

Therefore, as robustness check, we estimate the SDF-STE model in Eq.(3)-(4) considering different kinds of spatial weight matrices. Specifically, we substitute the dense inverse distance spatial weight matrix  $W$  used until now with different inverse distance spatial weight matrices truncated at a 200, 100, 50, and 30 kilometres radius around the  $i^{th}$  observation or considering the 400, 250, 100, 50, and 30 nearest neighbours. Descriptive statistics for these different spatial weight matrices are shown in Table D1 in Appendix D.

The results presented in Appendix D (Table D2 and Table D3) show that the  $\beta$  and the  $\phi$  estimates are robust to different changes of  $W$ , as well as the estimates associated with the two variance parameters  $\sigma^2$  and  $\lambda$ . As expected, only the spatial parameters  $\rho$ ,  $\theta$ , and  $\delta$  are affected by different choices of the spatial weight matrix. Specifically, the degree of global spatial dependence captured through  $\rho$  tends to decrease as the number of neighbours decreases, passing from a maximum value of 0.351 using a dense inverse distance  $W$ , to 0.132 and 0.131 considering a truncation point at 30km or the 30 nearest neighbours, respectively. This is due to a reduction of the  $w_{ij}$  values used to weight the spatial units, as shown in the first column of Table D1.

**Table 4:** Marginal Effects: Sensitivity to the Choice of  $W$

		W	W200t	W100t	W50t	W30t	W400n	W250n	W100n	W50n	W30n
L	Direct	0.743***	0.738***	0.738***	0.738***	0.738***	0.738***	0.738***	0.738***	0.739***	0.738***
	Indirect	0.071***	0.056***	0.048***	0.043***	0.038***	0.054***	0.049***	0.039***	0.038***	0.034***
	Total	0.814***	0.794***	0.079***	0.782***	0.776***	0.792***	0.787***	0.777***	0.777***	0.772***
K	Direct	0.136***	0.136***	0.148***	0.148***	0.147***	0.148***	0.148***	0.147***	0.148***	0.147***
	Indirect	0.082***	0.056***	0.048***	0.036***	0.023***	0.050***	0.046***	0.037***	0.030***	0.028***
	Total	0.218***	0.192***	0.196***	0.184***	0.170***	0.198***	0.194***	0.184***	0.178***	0.175***
t	Direct	0.006***	0.007***	0.007***	0.008***	0.009***	0.008***	0.008***	0.008***	0.008***	0.008***
	Indirect	0.004***	0.003***	0.002***	0.002***	0.001***	0.002***	0.002***	0.002***	0.002***	0.001***
	Total	0.010***	0.010***	0.009***	0.010***	0.010***	0.010***	0.010***	0.010***	0.010***	0.009***
Hum	Direct	-0.647***	-0.646***	-0.646***	-0.645***	-0.645***	-0.0646***	-0.645***	-0.645***	-0.645***	-0.645***
	Indirect	-0.103***	-0.063***	-0.065***	-0.066***	-0.056***	-0.069***	-0.070***	-0.066***	-0.063***	-0.061***
	Total	-0.750***	-0.709***	-0.711***	-0.710***	-0.700***	-0.715***	-0.715***	-0.711***	-0.708***	-0.706***
Int	Direct	-0.116***	-0.117***	-0.117***	-0.116***	-0.116***	-0.116***	-0.116***	-0.116***	-0.117***	-0.117***
	Indirect	-0.171***	-0.120***	-0.106***	-0.085***	-0.053***	-0.099***	-0.097***	-0.094***	-0.072***	-0.061***
	Total	-0.287***	-0.236***	-0.222***	-0.202***	-0.160***	-0.216***	-0.213***	-0.210***	-0.189***	-0.178***

Pat	Direct	-0.054***	-0.053***	-0.052***	-0.053***	-0.053***	-0.053***	-0.052***	-0.052***	-0.052***	-0.052***
	Indirect	0.122***	0.107***	0.073***	0.033	0.017	0.071***	0.044**	0.017	0.013	0.006
	Total	0.0688**	0.054*	0.021	-0.019	-0.036	0.018	-0.008	-0.035	-0.039	-0.046*
Trad	Direct	-0.037***	-0.038***	-0.038***	-0.039***	-0.038***	-0.038***	-0.038***	-0.039***	-0.039***	-0.038***
	Indirect	0.039*	0.031	0.034	0.016	0.004	0.047*	0.045*	0.025	0.031	0.03
	Total	0.002	-0.007	-0.04	-0.023	-0.034	0.009	0.007	-0.013	-0.007	-0.008
Size	Direct	-0.037***	-0.037***	-0.038***	-0.039***	-0.039***	-0.037***	-0.038***	-0.038***	-0.038***	-0.039***
	Indirect	-0.040***	-0.048***	-0.035**	-0.022*	-0.023**	-0.034***	-0.032***	-0.028**	-0.031***	-0.025**
	Total	-0.077***	-0.084***	-0.073***	-0.061***	-0.062***	-0.071***	-0.070***	-0.067***	-0.069***	-0.064***

\*\*\* :  $pvalue \leq 0.01$ ; \*\* :  $pvalue \leq 0.05$ ; \* :  $pvalue \leq 0.10$

Note: We define W200t, W100t, W50t, and W30t as inverse distance spatial weight matrices truncated at 200, 100, 50, and 30 kilometres respectively, while W400n, W250n, W100n, W50n, and W30n indicate inverse distance spatial weight matrices considering only the 400, 250, 100, 50, and 30 nearest neighbours, respectively.

To interpret the estimated coefficients in a meaningful way, direct, indirect, and total marginal effects are computed, as shown in Table 4. Results in Table 4 indicate that considering a dense inverse distance spatial weight matrix  $W$  leads to a slight overestimation of the indirect effects compared to the case when a sparse weight matrix is used. While the direct effect of labour and capital is quite stable across the different trials, the indirect effects of  $L$  and  $K$  pass from a maximum value of 0.071 and 0.082 to 0.034 and 0.023, respectively but they remain positive and significant. Similarly, the direct effects associated with  $Hum$ ,  $Int$ ,  $Pat$ ,  $Trad$  and  $Size$  are constant even if the spatial weight matrix changes while the indirect effects decrease as the number of considered neighbours is reduced.

Considering the indirect effects of the determinants of hotels' inefficiency, while indirect effects associated with human capital, intangible capital investments, and firm size are always negative and significant across all the different trials, the indirect effects of intellectual properties are only significant considering a truncation point higher than 100 km or more than 250 nearest neighbours for patents and using a dense spatial weight matrix or more than 250 neighbours for trademarks. Hence, patents and trademarks seem to generate negative spillovers and to effectively protect new products and ideas only considering high distances,

while the indirect effect resulting from having neighbours that have registered trademarks and/or patents is not significant for hotels that are very close to the innovative firm. Thus, negative spatial spillovers generated from registered patents or trademarks apply at the global level but they are not significant at a local level. Indeed, at the local level, positive externalities due to interpersonal contact and shared ideas at meetings and events can occur, cancelling the blocking function of patents and trademarks.

To sum up, results from this robustness check indicate that while the estimated direct effects are stable across the different specifications of  $W$ , the indirect effects tend to rise in magnitude as the geographical distance to identify neighbouring units increases. Thus, spatial spillover effects occurring in the Italian accommodation sector tend to cumulate across space, in line with the results of Cainelli and Ganau (2018) for the Italian manufacturing industry. Hence, it would be more effective for policy makers to develop plans aimed at fostering the innovativeness of the whole sector at the national level without focusing on single local areas to entirely exploit the existing spatial interactions.

#### ***5.4 Robustness Check***

Besides spatial individual heterogeneity, unobserved individual-specific effects such as entrepreneurial or managerial skills are likely relevant for hotel performance and the productive outcome of hotels may be endogenously related to the input variables or to the inefficiency determinants. Moreover, it is quite plausible that hotels with good prospects can decide to locate in areas close to competitors to benefit from local advantages, generating endogeneity issues due to omitted variables. However, to date, in stochastic frontier literature, there are no current available methods dealing together with spatial heterogeneity, individual heterogeneity and endogeneity. Indeed, recent advancements in stochastic frontier models literature have primarily focused on two different directions: (i) introducing some spatial components in order to consider cross-sectional spatial dependence (Glass et. al, 2016; Tsukamoto, 2019; Orea and Alvarez, 2019; Galli, 2021); (ii) controlling for non-spatial individual heterogeneity or for possible sources of endogeneity (Greene 2005a; 2005b; Wang and Ho, 2010; Amsler et. al, 2016; Belotti and Ilardi, 2018; Kutlu et al., 2019; Tsionas and Mallick, 2019). Up to now, to our knowledge, there are no contributions controlling for both spatial and non-spatial individual heterogeneity, while the only work developing a SF model considering both a spatial autoregressive term and endogeneity due to correlation of the inefficiency term and the two-sided error term is Kutlu (2020). Here, positioning in the first strand of literature, the SDF-STE model was estimated with the main aim of investigating the



spatial dimension of our phenomenon; the extension of the proposed model to include endogeneity and heterogeneous issues is left to future research.

However, given the relevant role of both individual effects and endogeneity issues, in this section, we compare our results to other SF approaches that allow considering individual specific effects or that control for possible endogeneity to test the robustness of our baseline estimates. Specifically, when dealing with individual heterogeneity we compare our non-spatial results corresponding to the Battese and Coelli (1995) specification with those of the non-spatial true fixed effect stochastic frontier model introduced by Greene (2005a) because at the moment, there is no available spatial SF model controlling for individual heterogeneity. On the other hand, following Castiglione (2014) and De Vries and Koetter (2011), we partially attempt to control for the presence of endogeneity by introducing in our SDF-STE model lagged input variables and lagged determinants of inefficiency. Specifically, we model the productive outcome of period  $t$  as a function of labour and capital at time  $t-1$  and of the 1-year lag of intangible capital and human capital investments. Furthermore, we also make a second robustness check considering a two-year lag. Starting from individual effects, as shown in Table E1 of Appendix E, our non-spatial estimates are robust to the different model specification and thus, unobserved individual heterogeneity has a negligible impact on the estimation results. Concerning endogeneity, the results shown in Table E2 of Appendix E confirm the robustness of our estimates to simultaneity issues related to the input variables and to the determinants of inefficiency. Finally, in Table E3 of Appendix E, we show the results of the final robustness check controlling for both individual fixed effects and for endogeneity using lagged variables starting from the non-spatial specification proposed by Battese and Coelli (1995). Overall, the results are in line with our baseline estimates, showing that distortions arising from individual unobserved effects and endogeneity are small. Thus, according with Rice et al. (2006), Koo and Lall (2007), Ellison et al. (2010), and Drucker (2012), we do not expect that endogeneity issues or individual heterogeneity could have a relevant and distortive impact on our findings.

## **6 Conclusion**

Scholars have widely acknowledged the relevance of spatial interactions in affecting the productivity level of hotels belonging to tourism clusters. However, to date, no studies have yet clearly identified the different typologies, the magnitude, and the sources of these spatial effects. Thus, using georeferenced data and taking advantage of the SDF-STE model (Galli, 2021), in this paper we provide new insights on the spatial spillover effects related to the

determinants of hotels' innovative activity. To the best of our knowledge, this is the first study aiming at investigating the role of both internal and external innovation in influencing the productivity and efficiency level of tourism facilities at a micro-economic level using appropriate spatial econometric techniques.

Results from this analysis indicate that the Italian accommodation sector is a labour-intensive sector with a high exploitation of internal human resources rather than investments in intangible capital. Therefore, to achieve higher profitability levels, hotels, from an internal point of view, rely more on innovation related to human capital, knowledgeable and skilled workers, and improved service quality than on product innovation generated by investments in innovative activities such as R&D and ITC (Q1). Considering spillover effects, we find that the innovative activity performed by neighbours significantly spreads across space (Q2) generating both agglomeration and competition effects (Q3). In particular, our results show that having nearby hotels that invest in labour and capital positively affects the level of productivity of neighbouring hotels. Similarly, having neighbours who invest in human capital has a positive effect on hotels' efficiency level as well as having neighbours making intense innovative activity. In particular, we detect a greater positive indirect effect of intangible investments on firm efficiency compared to the direct one, meaning that spillover effects generated by highly innovative hotels have a stronger cumulative impact on neighbouring firms than on the innovator itself. Therefore, despite being a labour-intensive sector, investments in intangible capital by a few innovative hotels contribute to the development of the whole Italian accommodation sector. In addition, having bigger hotels as peers has a positive effect on nearby firms. Conversely, registered patents and/or trademarks negatively affect neighbours thanks to a strong protection and blocking function. Thus, different sources of innovation generate different spatial effects, both positive and negative.

Our findings have important implications both from a theoretical and a practical perspective. From a scientific point of view, we empirically confirm the key role played by innovation in the hotel sector as a promoter of competitive advantage in tourism destinations. Moreover, in line with the industrial agglomeration theory, this paper extends to the Italian accommodation industry previous findings on other industrial sectors regarding the relevance of firms' location choices and of spatial interactions in influencing the level of competitiveness of neighbouring units. New insights from this study on spatial patterns affecting hotels' performance concern (i) the evidence of significant spatial effects both at the global and local level; (ii) the existence of different spillovers in terms of magnitude and signs resulting from the different sources of internal innovation. From a practical perspective, insights from this

analysis can be useful both for accommodation managers to improve their production processes by innovating and creating hotel networks and alliances, and for policy makers to design place-based policies supporting hotels' innovative activity and spatial interactions across tourism firms. Public incentives for the tourism sector should be aimed at stimulating hotels' innovative activity due to its high association with hotel performance. Since innovation is still an underdeveloped activity in the accommodation industry, external push factors are fundamental to spurring product and process innovation in the hospitality industry. Efficient innovation policies in this sector should stimulate innovations that allow energy savings and the sustainable management of resources in order to pursue sustainability, help hospitality businesses to create synergies that help overcome the limitations deriving from the small size, and motivate hotels' managers to personalize their offer thanks to the possibility of profiling customers in an increasingly specific and detailed manner. In addition, other accessible innovations such as IT adoption, improvements in customer service and in administrative practices, architectural and infrastructural renovation, and collaboration with the other actors in the sector should be promoted. Policy makers should therefore encourage accommodation facilities to network and create a healthy competitive environment allowing the transmission of new knowledge and innovation. In particular, innovative activity performed by bigger hotels tends to spread out to all neighbouring small and medium-sized hotels that are more unwilling to innovate. Therefore, by reinforcing hotels' networking and cooperation, the few big innovators in the sector may act as role models and knowledge disseminators for all those small entrepreneurs who are the main providers of hospitality services. This diffusion mechanism can foster and sustain the growth of the tourism sector and consequently, of the whole Italian economy.

A further extension of this work could also consider the food and beverage sector, the cultural and recreational services sector, and the transport sector. Merging the data from all these sectors with the accommodation industry would allow for the assessment of both intra-sectors and inter-sector spatial effects. Moreover, it would be interesting to evaluate the impact of different specific territorial characteristics on the efficiency level of the Italian accommodation sector after having retrieved technical efficiency scores.

## References

Adam, I. and E. A Mensah (2013). "Perceived spatial agglomeration effects and hotel location choice". *Anatolia* 25(1), 49-60.

- Adams, J.D. and A.B. Jaffe (1996). "Bounding the Effects of R&D: An Investigation Using Matched Establishment-Firm Data". *RAND Journal of Economics* 27(4), 700-721.
- Adetutu, M., A.J. Glass, K. Kenjegalieva and R.C. Sickles (2015). "The effects of efficiency and TFP growth on pollution in Europe: A multistage spatial analysis". *Journal of Productivity Analysis* 43, 307-326.
- Aigner, D., C.K. Lovell and P. Schmidt (1977). "Formulation and estimation of stochastic frontier production function models". *Journal of Econometrics* 6(1), 21-37.
- Amsler, C., A. Prokhov and P. Schmidt (2016). "Endogeneity in stochastic frontier models". *Journal of Econometrics* 190(2), 280-288.
- Arbia, G. (1989). *Spatial data configuration in statistical analysis of regional economic and related problems*. Dordrecht: Kluwer.
- Audretsch, D.B. (1998). "Agglomeration and the location of innovative activity". *Oxford Review of Economic Policy* 14(2), 18-29.
- Backman, M., J. Klaesson and Ö. Öner (2017). "Innovation in the hospitality industry - Firm or location?" *Tourism Economics* 23(8), 1591-1614.
- Barros, C.P. (2005). "Measuring efficiency in the hotel sector". *Annals of Tourism Research* 32(2), 456-477.
- Battese, G. (1997). "A note on the estimation of Cobb-Douglas production functions when some explanatory variables have zero values". *Journal of Agricultural Economics* 48(1-3), 250-252.
- Battese, G.E. and T.J. Coelli (1988). "Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data". *Journal of Economics* 38(3), 387-399.
- Battese, G.E. and T.J. Coelli (1995). "A model for technical inefficiency effects in a stochastic frontier production function for panel data". *Empirical Economics* 20(2), 325-332.
- Baum, J.A.C. and S.J. Mezas (1992). "Localized Competition and Organizational Failure in the Manhattan Hotel Industry, 1898-1990". *Administrative Science Quarterly* 37(4), 580-604.
- Belotti, F. and G. Ilardi (2018). Consistent inference in fixed-effects stochastic frontier models. *Journal of Econometrics* 202(2), 161-177.
- Benkovskis, K. (2018). "Misallocation, productivity and fragmentation of production: the case of Latvia". *Journal of Productivity Analysis* 49, 187-206.

- Bernini, C. and A. Guizzardi (2010). "Internal and locational factors affecting hotel industry efficiency: evidence from Italian business corporations". *Tourism Economics* 16(4), 883-913.
- Bernini, C. and A. Guizzardi (2016). "Internal features and agglomeration externalities for the hotels' competitiveness in Emilia Romagna." In Mariani, M.M., Buhalis, D., Czakon, W., Vitouladiti, O., (eds.), "Tourism Management, Marketing, and Development: Performance, Strategies, and Sustainability", New York: Palgrave.
- Binder, P. (2019). "A network perspective on organizational learning research in tourism and hospitality." *International Journal of Contemporary Hospitality Management* 31(7), 2602-2625.
- Brida, J.G., I. Cortes-Jimenez and M. Pulina (2016). "Has the tourism-led growth hypothesis been validated? A literature review". *Current Issues in Tourism* 19(5), 394-430.
- Cabrer-Borras, B. and G. Serrano-Domingo (2007). "Innovation and R&D spillover effect in Spanish regions: A spatial approach". *Research Policy* 36, 1357-1371.
- Camisón-Zornosa, C., R. Lapiedra-Alcamí, M. Segarra-Ciprés and M. Boronat-Navarro (2004). "A meta-analysis of innovation and organizational size". *Organization Studies* 25(3), 331-361
- Cainelli, G. and R. Ganau (2018). "Distance-based agglomeration externalities and neighbouring firms' characteristics". *Regional Studies* 52(7), 922-933.
- Carlino, G. and W.R. Kerr (2015). Chapter 6 - Agglomeration and Innovation. In Gilles Duranton, J. V. H. and Strange, W. C., Eds.; Elsevier: Amsterdam, The Netherlands, Volume 5, pp. 349-404.
- Castiglione, C. and D. Infante (2014). "ICTs and time-span in technical efficiency gains. A stochastic frontier approach over a panel of Italian manufacturing firms". *Economic Modelling* 41, 55-65.
- Chang, C. L., T. Khamkaew and M. McAleer (2012). "IV estimation of a panel threshold model of tourism specialization and economic development". *Tourism Economics* 18(1), 5-41.
- Chen, C. J., H.A. Shih and S.Y. Yang (2009). The role of intellectual capital in knowledge transfer. *IEEE Transactions on Engineering Management* 56(3), 402-411.
- Chyi, Y., Y. Lai and W. Liu (2012). "Knowledge spillovers and firm performance in the high-technology industrial cluster". *Research Policy* 41, 556-564.

- CDP (2019). Innovazione e hospitality: quali leve per la competitività del turismo in Italia?.  
[https://www.cdp.it/resources/cms/documents/2019-12-12\\_CDP\\_Ricerca\\_Innovazione%20del%20turismo.pdf](https://www.cdp.it/resources/cms/documents/2019-12-12_CDP_Ricerca_Innovazione%20del%20turismo.pdf)
- Cho, S., R.H. Woods, S. Jang and M. Erdem (2006). "Measuring the impact of human resource management practices on hospitality firms' performances". *Hospitality Management* 25, 262-277.
- Croes, R. (2013). Tourism specialization and economic output in small islands. *Tourism Review* 68(4), 34-48.
- Croes, R., J. Ridderstaat, M. Bak and P. Zientara (2021). "Tourism specialization, economic growth, human development and transition economies: The case of Poland". *Tourism Management* 82, 104181.
- De Vries, G.J. and M. Koetter (2011). "ICT Adoption and Heterogeneity in Production Technologies: Evidence for Chilean Retailers". *Oxford Bulletin of Economics and Statistics* 73(4), 0305-9049.
- Decelle, X. (2006). "A dynamic conceptual approach to innovation in tourism". Innovation and tourism policy. OECD. OECD Publishing, 85-105.
- Delgado, M., M.E. Porter and S. Stern (2014). "Clusters, convergence, and economic performance". *Research Policy* 43, 1785-1799.
- Drucker, J. and E. Feser (2012). "Regional industrial structure and agglomeration economies: An analysis of productivity in three manufacturing industries". *Regional Science and Urban Economics* 42, 1-14.
- Ellison, G., E.L. Glaeser and W.R. Kerr (2010). "What causes industry agglomeration? Evidence from coagglomeration patterns". *The American Economic Review* 100(3), 1195-1213.
- Galli, F. (2021). "A Spatial Durbin Stochastic Frontier Model Introducing Spillover Effects in the Determinants of Firms' Efficiency". Available at SSRN:  
<https://ssrn.com/abstract=3924100>
- Garcia-Ayuso, M., I. Moreno-Campos and G. Sierra-Molina (2000). "Fundamental analysis and human capital: Empirical evidence on the relationship between the quality of human resources and fundamental accounting variables." *Journal of Human Resource Costing & Accounting* 5(1), 45-57.
- Glass, A.J., K. Kenjegaliev and R.C. Sickles (2016). "A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers". *Journal of Econometrics* 190(2), 289-300.

- Gotsch, M. and C. Hipp (2012). "Measurement of Innovation Activities in the Knowledge Intensive Services Industry: A Trademark Approach". *The Service Industries Journal* 32(13), 2167–2184.
- Greene, W. (2005a). "Fixed and Random Effects in Stochastic Frontier Models". *Journal of Productivity Analysis* 23, 7-32.
- Greene, W. (2005b). "Reconsidering heterogeneity in panel data estimators of the stochastic frontier model". *Journal of Econometrics* 126, 269-303.
- Griffith, R., R. Harrison and J. van Reenen (2006). "How Special is the Special Relationship? Using the Impact of U.S. R&D Spillovers on U.K. Firms as a Test of Technology Sourcing". *American Economic Review* 96(5), 1859–1875.
- Hameed, W.U., Q.A. Nisar and H. Wu (2021). "Relationships between external knowledge, internal innovation, firms' open innovation performance, service innovation and business performance in the Pakistani hotel industry". *International Journal of Hospitality Management* 92, 102745.
- Haschka, R. and H. Herwartz (2020). "Innovation efficiency in European high-tech industries: Evidence from a Bayesian stochastic frontier approach". *Research Policy* 49, 104054.
- Hjalager, A.M. (2002). "Repairing innovation defectiveness in tourism". *Tourism Management* 23, 465-474.
- Hjalager, A.M. (2010). "A review of innovation research in tourism". *Tourism Management* 31, 1-12.
- Hsieh, C.T. and P.J. Klenow (2009). "Misallocation and manufacturing TFP in China and India". *The Quarterly Journal of Economics* 124(4), 1403-1448.
- Jacob, M. and J.L. Groizard (2007). "Technology transfer and multinationals: the case of Balearic hotel chains' investments in two developing economies". *Tourism Management* 28, 976–992.
- Jacob, M., C. Florido and E. Aguiló (2010). "Environmental innovation as a competitiveness factor in the Balearic Islands". *Tourism Economics* 16(3), 755–764.
- Jackson, J. and P. Murphy (2002). "Tourism destinations as clusters: Analytical experiences from the New World". *Tourism and Hospitality Research* 4(1), 36–52.
- Jackson, J. and P. Murphy (2006). "Clusters in regional tourism An Australian Case". *Annals of Tourism Research* 33(4), 1018–1035.
- Jaremen, D.E. (2016). "Advantages from ICT usage in hotel industry". *Czech Journal of Social Sciences, Business and Economics* 5(3), 6-17.

- Jiménez-Zarco, A.I., M.P. Martínez-Ruiz and A. Izquierdo-Yusta (2011). "Key service innovation drivers in the tourism sector: empirical evidence and managerial implications". *Service Business* 5, 339–360.
- Lajili, K. and D. Zeghal (2006). "Market performance impacts of human capital disclosures". *Journal of Accounting and Public Policy* 25(2), 171–194.
- Kim, Y.R., A.M. Williams, S. Park and J. Li Chen (2021). "Spatial spillovers of agglomeration economies and productivity in the tourism industry: The case of the UK". *Tourism Management* (82), 104201.
- Koch, L. and M. Simmler (2020). "How important are local knowledge spillovers of public R&D and what drives them?". *Research Policy* 49, 104009.
- Koo, J. and S.V., Lall (2007). "New economic geography: real or hype?" *International Regional Science Review* 30(1), 3–19.
- Kutlu, L. (2020). "Spatial Stochastic Frontier Model with Endogenous Spillovers". Available at SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3626688](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3626688).
- Kutlu, L., K.C., Tran and M.G. Tsionas (2019). "A time-varying true individual effects model with endogenous regressors" *Journal of Econometrics* 211(2), 539-559.
- La Porta, R. and A. Shleifer (2008). "The Unofficial Economy and Economic Development". *Brookings Papers on Economic Activity* 39(2), 275-363.
- Le, Q.C. and R. Pomfret (2011). "Technology spillovers from foreign direct investment in Vietnam: horizontal or vertical spillovers?". *Journal of the Asia Pacific Economy* 16(2), 183-201.
- LeSage, J.P. and R.K Pace (2009). *Introduction to Spatial Econometrics*. Boca-Raton: Taylor & Francis.
- Lev, B., and A. Schwartz (1971). "On the use of the economic concept of human capital in financial statements." *The Accounting Review* 46(1), 103–112.
- Levin, R.C. and P. Reiss (1988). "Cost-reducing and demand-creating R&D with spillovers". *Rand Journal of Economics* 19, 538–556.
- Marco-Lajara, B., E. Claver-Cortés, M. Úbeda García, F. García-Lillo and P.C. Zaragoza-Sáez (2019). "The role of internal knowledge generation and external knowledge acquisition in tourist districts". *Journal of Business Research* 101, 767–776.
- Marco-Lajara, B., P.C. Zaragoza-Sáez, E. Claver-Cortés and M. Úbeda García (2016). "Knowledge sources, agglomeration, and hotel performance". *Journal of Business Research* 69, 4856–4861.



- Marshall, A. (1890). *Principles of Economics*. London: Macmillan.
- Martin, R. L. and M.C. Moldoveanu (2003). "Capital versus talent: The battle that's reshaping business". *Harvard Business Review* 81(7), 36–41.
- Medina-Munoz, D., R.D. Medina-Munoz and A. Zuniga-Collazos (2013). "Tourism and innovation in China and Spain: a review of innovation research on tourism". *Tourism Economics* 19(2), 319-337.
- Montresor S. and A. Vezzani (2016). "Intangible investments and innovation propensity: Evidence from the Innobarometer 2013". *Industry and Innovation* 23(4), 331-352.
- Nelson A.J., (2009). "Measuring knowledge spillovers: What patents, licenses and publications reveal about innovation diffusion". *Research Policy* 38, 994-1005.
- O'Mahony, M. and M. Vecchi (2009). "R&D, knowledge spillovers and company productive performance". *Research Policy* 38, 35-44.
- Orea, L. and I.C. Alvarez (2019). "A new stochastic frontier model with cross-sectional effects in both noise and inefficiency terms". *Journal of Econometrics* 213, 556-577.
- Orfila-Sintes F., R. Crespi-Caldera and E. Martinez-Ros (2005). "Innovation activity in the hotel industry: Evidence from Balearic Islands". *Tourism Management* 26(6), 851-865.
- Ottensbacher, M.C. (2007). "Innovation management in the hospitality industry: different strategies for achieving success". *Journal of Hospitality and Tourism Research* 31, 431–454.
- Peiró-Signes, A., M. Segarra-Oña, L. Miret-Pastor and R. Verma (2015). "The Effect of Tourism Clusters on U.S. Hotel Performance". *Cornell Hospitality Quarterly* 56(2), 155–167.
- Poon, A. (1993). "Tourism, technology and competitive strategies". Wallingford: CAB International.
- Poon, A. (1994). "The 'new tourism' revolution". *Tourism Management* 15(2), 91–92.
- Pulakos, E. D., D.W. Dorsey and W.C. Borman (2003). "Hiring for knowledge-based competition". In S. E. Jackson, A. DeNisi, & M. Hitt (Eds.), *Managing knowledge for sustained competitive advantage* (pp. 155–177). San Francisco, CA: John Wiley and Sons.
- Rhee, K.H. and H.K. Pyo (2010). "Financial crisis and relative productivity dynamics in Korea: evidence from firm-level data (1992-2003)". *Journal of Productivity Analysis* 34, 111-131.

- Rice, P., A.J. Venables and E. Patacchini (2006). "Spatial determinants of productivity: analysis for the regions of Great Britain." *Regional Science and Urban Economics* 36(6), 727-752.
- Ridderstaat, J., R. Croes and P. Nijkamp (2016). "The tourism development-quality of life nexus in a small island destination". *Journal of Travel Research*, 55(1), 79-94.
- Romero, M.P. and J.A. Molina (2013). "Tourism and Economic Growth: A Review of Empirical Literature". *Tourism Management Perspectives* 8, 28-41.
- Sari, D.W., N.A. Khalifa and S. Suyanto (2016). "The spillover effects of foreign direct investment on the firms' productivity performances". *Journal of Productivity Analysis* 46, 199-233.
- Shaw, G. and A. Williams (2009). "Knowledge transfer and management in tourism organisations: An emerging research agenda". *Tourism Management* 30, 325-335.
- Schubert, S.F., J.G. Brida and W.A. Risso(2011). "The impacts of international tourism demand on economic growth of small economies dependent on tourism". *Tourism Management* 32(2), 377-385.
- Soteriades, M., C. Aivalis and D. Varvaressos (2004). "E-Marketing and E-Commerce in The Tourism Industry: A Framework to Develop and Implement Business Initiatives". *Tourism Today* 4, 1-18.
- Spence, M. (1984). "Cost Reduction, Competition, and Industry Performance". *Econometrica* 52, 101-121.
- Stojcic, N., K. Vojvodic and N. Butigan (2019). "Service Innovation and Firm Performance in the Hospitality Industry: Evidence from Tourism Driven Economy". *Tourism in Southern and Eastern Europe* 5, 679-688.
- Succurro, M. and F. Boffa (2018). "Patenting patterns in the tourism industry: Evidence from Italy". *International Journal of Tourism Research*, 20, 538-541.
- Sundbo, J., F. Orfila-Sintes and F. Sørensen (2007). "The innovative behaviour of tourism firms – Comparative studies of Denmark and Spain". *Research Policy* 36, 88-106.
- Sydler, R., S. Haefliger and R. Pruks (2014). "Measuring intellectual capital with financial figures: Can we predict firm profitability?" *European Management Journal* 32, 244-259.
- Syverson, C. (2011). "What determines productivity?" *Journal of Economic Literature* 49(2), 326-365.
- Tappeiner, G., C. Hauser and J. Walde (2008). "Regional knowledge spillovers: Fact or artifact?". *Research Policy* 37, 861-874.

- Tsionas, M. and S.K. Mallick (2019). "A Bayesian semiparametric approach to stochastic frontiers and productivity". *European Journal of Operational Research* 274, 391-402.
- Tsukamoto, T. (2019). "A spatial autoregressive stochastic frontier model for panel data incorporating a model of technical inefficiency". *Japan and the World Economy* 50, 66-77.
- Wakelin, K. (1998). "Innovation and export behaviour at the firm level". *Research Policy* 26(7-8), 829-841.
- Wang, H.-J. and C.-W. Ho (2010). Estimating fixed-effects panel stochastic frontier models by model transformation. *Journal of Econometrics* 157(2), 286-296.
- Weidenfeld, A., A.M. Williams and R.W. Bultler (2010). "Knowledge transfer and innovation across attractions". *Annals of Tourism Research* 37(3), 604-626.
- Williams, A.M., I. Rodriguez Sanchez and V. Skokic (2021). "Innovation, Risk and Uncertainty: A Study of Tourism Entrepreneurs". *Journal of Travel Research* 60(2), 293-311.
- Wyatt, A., and H. Frick (2010). "Accounting for investments in human capital: A review". *Australian Accounting Review* 20(3), 199-220.
- Yang, Y. (2012). "Agglomeration density and tourism development in China: An empirical research based on dynamic panel data model". *Tourism Management* 33, 1347-1359.
- Zhang, C., H. Xiao, D. Gursoy and Y. Rao (2015). "Tacit knowledge spillover and sustainability in destination development". *Journal of Sustainable Tourism* 23(7), 1029-1048.

## **Supplementary Materials**

## Appendix A

### Data Cleaning Procedure

The data used for the analysis were collected from the AIDA - Bureau van Dijk database. The AIDA databank is largely used in empirical research because of the high coverage of both the firms observed within sectors and balance sheet information. In our analysis the sample coverage is on average 12.7% in terms of firms belonging to the Italian accommodation sector, reaching 12.65% for firms with less than 200 employees and 57.89% for firms with more than 200 employees. These rates are much higher than the coverage of alternative surveys such as the ISTAT survey on Income Accounts of Enterprises, whose sample coverage is on average about 2% for all Italian firms with less than 250 employees and 79% for all Italian firms over 250 employees. Moreover, only in the AIDA databank, the geographical localisation is provided, allowing us to implement our spatial analysis. All these advantages motivated us to use the AIDA databank for our analysis. Specifically, we downloaded all data referring to firms belonging to the ATECO55 sector, that is the Italian accommodation sector, in the time period 2011-2019.

Starting from a sample of more than 20,000 individual observations available yearly, we ended up with a balanced panel of 5409 firms for each year due to a necessary cleaning procedure. Firstly, we dropped all the observations having missing values for the value added and for the number of employees in all the years covered by the analysis and we also dropped all the observations having negative values for the value added in at least one year. Then, we dropped all firms having non-active legal status, ending up with a sample of 14,241 firms. Next, we interpolated the missing values in the variables value added, number of employees, fixed capital, immaterial capital and personnel costs and then we dropped all the observations still reporting at least one missing value in at least one year after the interpolation procedure. The percentage of interpolated values is 12.4% in 2019, 5.9% in 2018, 8.7% in 2017, 11.2% in 2016, 13.6% in 2015, 15.1% in 2014, 16.7% in 2013, 17.3% in 2012, 17.7% in 2011. The mean of the whole period is 13.2%. Afterward, we dropped all the observations having value added, fixed capital and personnel costs less than one thousand and number of employees less than one to avoid generating missing values computing logarithms. In the end, we obtained a final cleaned sample consisting in 5,409 observations yearly and in total 48681 overall. Finally, starting from the addresses provided by the AIDA database, we geolocated each observation

using the R package “ggmap” which exploits the Google Geocoding API service of the Google Cloud Platform Console to find the latitude and longitude of each hotel.

In Table A1 we show the estimation results of the SDF-STE model for the year 2019 ( $T=1$ ) comparing our final sample ( $N=5409$ ) and all observations available for the year 2019 after having dropped all the missing values and without interpolating ( $N=7740$ ). The comparison is performed considering a one-year sample because to date, no spatial estimators for unbalanced panels are available for spatial stochastic frontier models. Indeed, with unbalanced panel, the spatial weight matrix  $W$  changes in each time period, leading to further issues to be addressed in the estimation procedure. The results shown in Table A1 indicate that our results are robust to different samples and to the presence of interpolated values, dissipating possible matters in our data due to the necessary cleaning procedure.

**Table A1.** Comparison between final sample and overall sample, year 2019.

	N=5409		N=7740	
	<i>Coeff.</i>	<i>t-stats</i>	<i>Coeff.</i>	<i>t-stats</i>
$\beta_0$	5.22***	28.40	5.35***	19.33
$\beta_L$	0.71***	30.05	0.73***	49.94
$\beta_K$	0.08***	6.08	0.10***	8.93
$\beta_{LL}$	0.07***	17.07	0.06***	17.03
$\beta_{KK}$	0.01***	9.69	0.01***	9.55
$\beta_{LK}$	-0.04***	-11.14	-0.04***	-13.64
$\rho$	0.31***	5.24	0.30***	5.32
$\theta_L$	-0.25***	-3.39	-0.29***	-5.05
$\theta_K$	0.04*	1.55	0.05**	2.12
$\phi_0$	5.37***	30.78	5.41***	19.29
$\phi_{hum}$	-0.70***	-66.08	-0.73***	-87.33
$\phi_{Int}$	-0.10***	-8.33	-0.10***	-10.79
$\phi_{pat}$	-0.05***	-3.01	-0.03**	-2.11
$\phi_{trad}$	-0.03*	-1.56	-0.01	-0.30
$\phi_{size}$	-0.04***	-4.07	-0.03***	-2.95
$\phi_{dsize}$	0.06*	1.47	0.02	0.65
$\delta_{hum}$	0.06	0.82	0.11*	1.51
$\delta_{Int}$	-0.10*	-1.26	-0.04	-0.67
$\delta_{pat}$	0.13	1.23	0.08	0.04
$\delta_{trad}$	0.03	0.23	-0.02	-0.11
$\delta_{size}$	0.06	1.01	0.07	1.28
$\delta_{dsize}$	-0.18	-0.62	-0.03	-0.13
Destination dummies	yes	-	yes	-
$\sigma^2$	0.21	-	0.21	-
$\lambda$	0.98	-	0.94	-

\*\*\*: *pvalue* ≤ 0.01; \*\*: *pvalue* ≤ 0.05; \*: *pvalue* ≤ 0.10; omit=omitted

## **Appendix B**

### **Comparison between sample and population**

Table B1 and Table B2 compare our sample with the corresponding population by class of employees and macro area. In particular, the former shows the coverage of our sample referring to the number of firms while the latter concerns the total number of employees. Population data were retrieved from the Industry and Services Census carried out by Istat in 2011. Table B1 shows that firms' coverage rate is lower for smaller hotels (on average 6.01% for hotels from 1 to 5 employees) while it increases considering bigger hotels, reaching the maximum value of 73.78% for hotels with 50-99 employees. Overall, the representativeness of our sample is good, covering the 12.69% of the ATECO 55 population. Considering the number of employees in 2011, the coverage rate of our sample is very good, reaching a value of 35.72% overall. As in the previous case, the coverage rate is lower for smaller hotels (15.11%) compared to bigger ones but, considering the total number of employees, the difference is less remarkable. Moreover, both for the number of firms and for the total number of employees, the coverage rate is higher for hotels located in the South of Italy and in the Islands while it is smaller for firms located in the North-East of Italy.

Examining hotels' coverage rate by Italian municipalities, Figure B1 shows that our sample is quite evenly distributed on the Italian territory. In particular, 2969 out of 7904 Italian municipalities in year 2018 do not have any tourist facility in their territory (undefined category). Considering the coverage rate of our sample, in 3528 municipalities of the remaining 4935 the coverage rate is smaller than 3%, while in 1407 municipalities it is higher than 3%.



**Table B1: Comparison between Sample and Population  
Number of Firms, Year 2011**

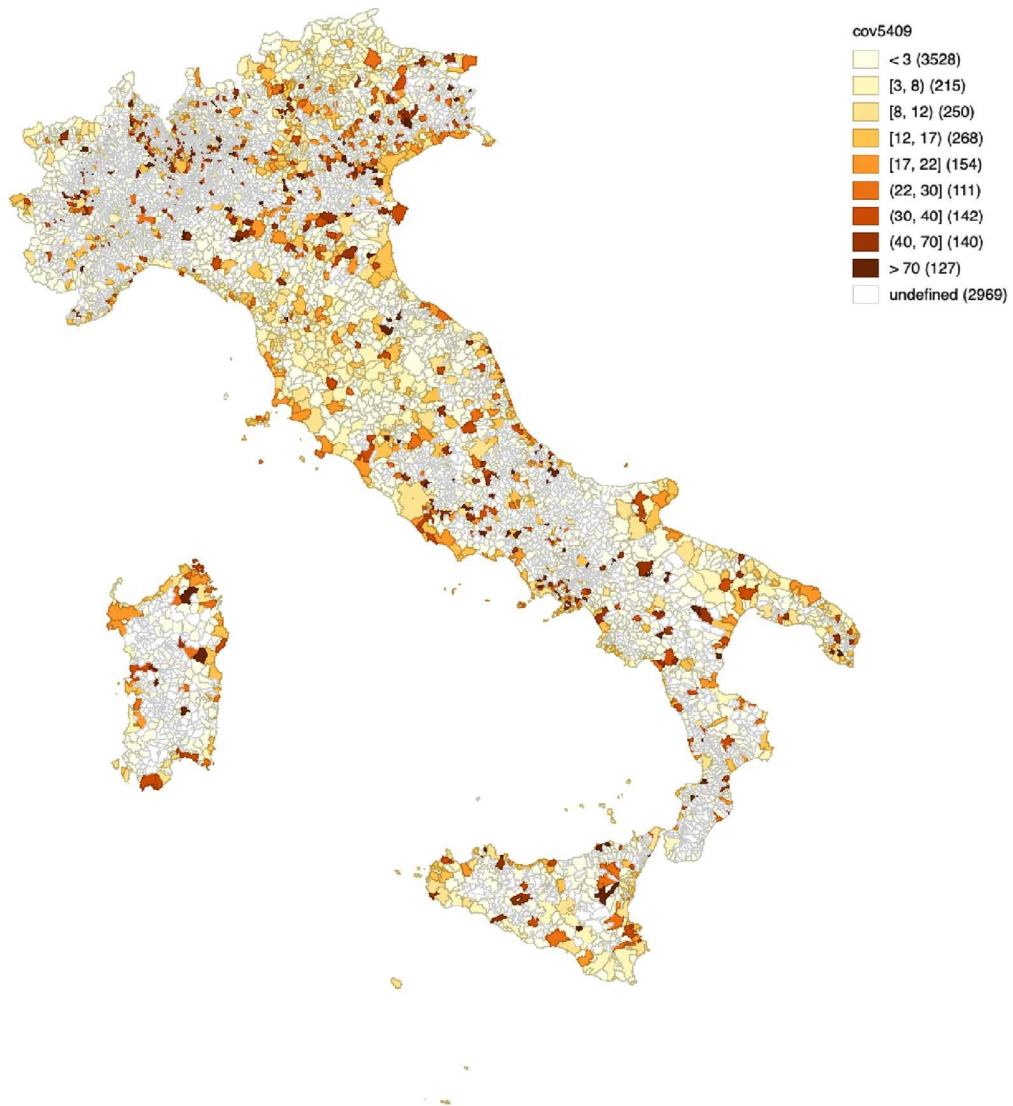
Macroarea		Class of Employees								Tot.
		1-5	6-9	10-15	16-19	20-49	50-99	100-199	200+	
North-West	<i>Pop.</i>	5396	916	579	157	236	35	17	14	7350
	<i>Sample</i>	362	189	226	74	129	29	12	9	1029
	<i>Cov.</i>	6.71	20.63	39.03	47.13	54.66	82.86	70.59	64.29	14.00
North-East	<i>Pop.</i>	11487	1520	1074	311	584	87	20	5	15088
	<i>Sample</i>	448	274	281	115	256	56	9	5	1442
	<i>Cov.</i>	3.90	18.03	26.16	36.98	43.84	64.37	45.00	100.00	9.56
Center	<i>Pop.</i>	8405	878	575	149	238	50	16	15	10326
	<i>Sample</i>	553	287	267	76	143	36	13	4	1378
	<i>Cov.</i>	6.58	32.69	46.43	51.01	60.08	72.00	81.25	26.67	13.34
South	<i>Pop.</i>	5464	546	334	114	207	36	8	2	6711
	<i>Sample</i>	402	185	162	87	151	34	8	2	1034
	<i>Cov.</i>	7.36	33.88	48.50	76.32	72.95	94.44	100.00	100.00	15.41
Islands	<i>Pop.</i>	2602	262	142	43	76	17	5	2	3149
	<i>Sample</i>	241	87	94	27	58	11	5	2	526
	<i>Cov.</i>	9.26	33.21	66.20	62.79	76.32	64.71	100.00	100.00	16.70
Tot.	<i>Pop.</i>	33354	4122	2704	774	1341	225	66	38	42624
	<i>Sample</i>	2006	1022	1030	379	737	166	47	22	5409
	<i>Cov.</i>	6.01	24.79	38.09	48.97	54.96	73.78	71.21	57.89	12.69

*Pop.*=Population; *Cov.*=Coverage

**Table B2: Comparison between Sample and Population  
Number of Employees, Year 2011**

Macroarea		Class of Employees								Tot.
		1-5	6-9	10-15	16-19	20-49	50-99	100-199	200+	
North-West	<i>Pop.</i>	11596	6603	6969	2723	6802	14913	2301	6391	58298
	<i>Sample</i>	4240	1403	2723	1282	3830	2001	1557	2365	19401
	<i>Cov.</i>	36.56	21.25	39.07	47.08	56.31	13.42	67.67	37.01	33.28
North-East	<i>Pop.</i>	22536	11123	13015	5386	16603	5779	2819	2190	79451
	<i>Sample</i>	828	2070	3405	1996	7589	3839	1217	2190	22280
	<i>Cov.</i>	3.67	18.61	26.16	37.06	45.71	66.43	43.17	100.00	28.04
Center	<i>Pop.</i>	16090	6383	6909	2556	6930	3219	2273	4866	49226
	<i>Sample</i>	2006	2146	3267	1319	4104	2267	1922	752	17914
	<i>Cov.</i>	12.47	33.62	47.29	51.60	59.22	70.43	84.56	15.45	36.39
South	<i>Pop.</i>	10116	3974	4050	1989	6010	2461	1012	613	30225
	<i>Sample</i>	2189	1382	1955	1497	4447	2282	1012	613	15822
	<i>Cov.</i>	21.64	34.78	48.27	75.26	73.99	92.73	100.00	100.00	52.35
Islands	<i>Pop.</i>	4619	1894	1714	742	2273	1171	589	922	13924
	<i>Sample</i>	551	638	1176	467	1705	804	589	922	7130
	<i>Cov.</i>	11.93	33.69	68.61	62.94	75.01	68.66	100.00	100.00	51.21
Tot.	<i>Pop.</i>	64957	29977	32657	13396	38618	27543	8994	14982	231124
	<i>Sample</i>	9814	7639	12526	6561	21675	11193	6297	6842	82547
	<i>Cov.</i>	15.11	25.48	38.36	48.98	56.13	40.64	70.01	45.67	35.72

*Pop.*=Population; *Cov.*=Coverage



**Figure B1.** Sample Coverage Map by Municipalities

## Appendix C

### Further details on the SDF-STE model

Appendix C contains more methodological insights on the assumptions, characteristics, and estimation of the spatial Durbin stochastic frontier model introducing spatial dependence in the determinants of firms' inefficiency (SDF-STE) by Galli (2021). The model specification is defined as in Eq.(1)-(4)

$$Y_{it} = X_{it}\beta + \rho \sum_{j=1}^N w_{ij} Y_{jt} + \sum_{j=1}^N w_{ij} X_{jt}\theta + v_{it} - u_{it} \quad (1)$$

$$v_{it} \sim N(0, \sigma_v^2) \quad (2)$$

$$u_{it} \sim N^+(\mu_{it}, \sigma_u^2) \quad (3)$$

$$\mu_{it} = Z_{it}\phi + \sum_{j=1}^N w_{ij} Z_{jt}\delta \quad (4)$$

where  $Y_{it}$  represents the output of firm  $i$  at time  $t$ ,  $X_{it}$  is a vector containing the  $k$  production inputs, and  $Z_{it}$  is a vector including  $m$  exogenous variables that are the determinants of firms' inefficiency level. The SDF-STE model also includes the spatial lag of the dependent variable, of the input variables and of the inefficiency determinants in order to consider the overall level of global spatial dependence and local spatial correlation related to the input variables and to the determinants of firms' inefficiency, respectively. Therefore,  $\rho$  captures the overall level of global spatial dependence while the  $\theta$  and  $\delta$  vector of parameters allow to measure the indirect effects related to the  $X$  and  $Z$  variables, respectively. As usual in spatial models, neighbouring units are identified through the spatial weight matrix  $W$  containing non-negative spatial weights  $w_{ij}$  for each pair of neighbours  $(i, j)$  and zero elements on the diagonal for  $i = j$ . Finally,  $v_{it}$  and  $u_{it}$  respectively represent the normally distributed random error term and the inefficiency error component that, in this framework, is commonly assumed to follow a half normal distribution since inefficiency can only take positive or at least zero values.

Assumptions on Eq.(1)-(4), following Elhorst (2010), include (i)  $(I_{NT} - \rho W)$  non-singular, where  $I_{NT}$  is the  $(NT \times NT)$  identity matrix; (ii) row and column sums of  $W$  and  $(I_{NT} - \rho W)^{-1}$ , before  $W$  is row-normalized, are uniformly bounded in absolute value as  $N$  goes to infinity

(Kelejian and Prucha, 1998; 1999). For a symmetric  $W$  the first assumption is always satisfied as long as the range of  $\rho$  is defined by  $\left(\frac{1}{\omega_{min}}, 1\right)$ , where  $\omega_{min}$  is the smallest real characteristic root of the spatial weight matrix  $W$  while the upper bound equals 1 for row-normalized  $W$ . Assumption (ii) limits the cross-sectional correlation, assuming that, when the distance separating two spatial units increases to infinity, it converges to zero. In particular, if  $W$  is a distance inverse spatial weight matrix, assumption (ii) can be guaranteed imposing a cut-off point  $d^*$  in  $W$  so that  $w_{ij} = 0$  if  $d_{ij} > d^*$ , while assumption (ii) is always satisfied if  $W$  is a binary contiguity matrix.

The SDF-STE model nests several existing spatial and non-spatial SF models. Imposing  $\delta = 0$  and  $\theta = 0$  the model reduces to the spatial autoregressive stochastic frontier model for panel data incorporating a model for technical inefficiency (SARF-TE) proposed by Tsukamoto (2019). If  $\delta = 0$  and  $\phi = 0$  the model becomes the spatial Durbin stochastic frontier model (SDF) introduced by Glass et al. (2016). Moreover, if  $\delta = 0$ ,  $\phi = 0$  and  $\theta = 0$  it coincides with the spatial autoregressive stochastic frontier model (SARF) by Glass et al. (2016). Imposing  $\delta = 0$ ,  $\rho = 0$  and  $\phi = 0$  it becomes the spatial stochastic frontier model introduced by Adetutu et al. (2015) that only includes the spatial lag of the exogenous variables (SLXF). Considering non-spatial SF model, if  $\delta = 0$ ,  $\theta = 0$  and  $\rho = 0$  the SDF-STE model reduces to the stochastic frontier production function with a model for technical inefficiency effects (SF-TE) proposed by Battese and Coelli (1995). Finally, considering  $\delta = 0$ ,  $\theta = 0$ ,  $\rho = 0$  and  $\phi = 0$  our model becomes the classical SF model by Aigner et al. (1977). Therefore, following an approach similar to Manski (1993) for spatial models, the SDF-STE model allows for various parametric restrictions, enabling a large set of modifications. Indeed, by implementing likelihood ratio tests and starting from our general specification, it is possible to select the model that best fits the data.

The likelihood function associated with the SDF-STE model can be calculated starting from the probability density functions of  $v_{it}$  and  $u_{it}$ . In particular, the former has a normal distribution with zero mean 0 and variance  $\sigma_v^2$  while the latter is distributed as a truncated normal random variable with mean  $\mu_{it}$  and variance  $\sigma_u^2$ . Therefore, the joint probability density function of  $v_{it}$  and  $u_{it}$ , assuming that  $v_{it}$  and  $u_{it}$  are independent, can be calculated as the product of their probability density functions  $f_v(v_{it})$  and  $f_u(u_{it})$ . Substituting  $v_{it} = \varepsilon_{it} + u_{it}$  in  $f_{uv}(u_{it}, v_{it})$ , starting from the relationship  $\varepsilon_{it} = v_{it} - u_{it}$ , it can be obtained the joint probability density function of  $\varepsilon_{it}$  and  $u_{it}$ . Afterwards, the probability density function of  $\varepsilon_{it}$  is obtained integrating out  $u_{it}$  from  $f_{u\varepsilon}(u_{it}, \varepsilon_{it})$ . Finally, starting from the joint probability

density function of  $\varepsilon$  obtained multiplying all the marginal distributions of  $\varepsilon_{it}$ , the probability density function of  $Y_{it}$  can be defined as the product of  $f_{\varepsilon}(\varepsilon)$  and of the determinant of the Jacobian of the transformation from  $\varepsilon_{it}$  to  $Y_{it}$ . Indeed, the endogeneity deriving from the inclusion of the spatial lag of the dependent variable has to be taken into account.

The final loglikelihood function, assuming that the panel is balanced, is given by

$$\begin{aligned} \mathcal{L}(\Theta; y) = & \log|I_{NT} - \rho W| - \frac{NT}{2} (\log \sigma^2 + \log 2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^N \sum_{t=1}^T (\mu_{it} + \varepsilon_{it})^2 \\ & - \sum_{i=1}^N \sum_{t=1}^T \left[ \log \Phi \left( \frac{\mu_{it}}{\sigma\sqrt{\lambda}} \right) - \log \Phi \left( \frac{\mu_{it}(1-\lambda) - \varepsilon_{it}\lambda}{\sigma\sqrt{\lambda(1-\lambda)}} \right) \right], \end{aligned} \quad (C1)$$

where  $\Theta$  represents the vector of all parameters,  $\Phi$  is the cumulative distribution function of the standard normal random variable and  $\mu_{it}$  and  $\varepsilon_{it}$  are defined as

$$\mu_{it} = Z_{it}\phi + \sum_{j=1}^N w_{ij}Z_{jt}\delta \quad (C2)$$

$$\varepsilon_{it} = Y_{it} - X_{it}\beta - \rho \sum_{j=1}^N w_{ij}Y_{jt} - \sum_{j=1}^N w_{ij}X_{jt}\theta. \quad (C3)$$

The parameter estimates can be obtained using a numerical maximization algorithm implemented in a standard software. Since the parameter space for an autoregressive process is  $(\frac{1}{\omega_{min}}, 1)$ , where  $\omega_{min}$  is the smallest eigenvalue of  $W$ , the autoregressive parameter  $\rho$  should be bounded to the previous interval. Moreover,  $\sigma^2$  should be positive and  $0 \leq \lambda \leq 1$ . Specifically, if  $\lambda$  equals zero the OLS model should be preferred to the SF function because the variance of the inefficiency term is zero and therefore, the determinants of firms' efficiency can be included in the frontier function. Conversely,  $\lambda$  increases until 1 if the inefficiency effects are likely to be highly significant. Finally, to make the algorithm work better, the first derivatives of the loglikelihood function with respect to the unknown parameters can be supplied to the program.

- Adetutu, M., A.J. Glass, K. Kenjegalieva and R.C. Sickles (2015). "The effects of efficiency and TFP growth on pollution in Europe: A multistage spatial analysis". *Journal of Productivity Analysis* 43, 307–326.
- Aigner, D., C.K. Lovell and P. Schmidt (1977). "Formulation and estimation of stochastic frontier production function models". *Journal of Econometrics* 6(1), 21–37.
- Battese, G.E. and T.J. Coelli (1995). "A model for technical inefficiency effects in a stochastic frontier production function for panel data". *Empirical Economics* 20(2), 325–332.
- Elhorst, J.P. (2010). Spatial panel data models. Fischer, M.M., Getis, A. (Eds.), *Handbook of Applied Spatial Analysis*. New York: Springer.
- Galli, F. (2021). "A Spatial Durbin Stochastic Frontier Model Introducing Spillover Effects in the Determinants of Firms' Efficiency". Available at SSRN: <https://ssrn.com/abstract=3924100>
- Glass, A.J., K. Kenjegalieva and R.C. Sickles (2016). "A spatial autoregressive stochastic frontier model for panel data with asymmetric efficiency spillovers". *Journal of Econometrics* 190(2), 289–300.
- Kelejian, H.H. and I.R. Prucha (1998). "A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbance". *The Journal of Real Estate Finance and Economics* 17, 99–121.
- Kelejian, H.H. and I.R. Prucha (1999). "A generalized moments estimator for the autoregressive parameter in a spatial model". *International Economic Review* 40(2), 509–533.
- Manski, C.F. (1993). "Identification of endogenous social effects: the reflection problem". *Review of Economic Studies* 60, 531–542.
- Tsukamoto, T. (2019). "A spatial autoregressive stochastic frontier model for panel data incorporating a model of technical inefficiency". *Japan and the World Economy* 50, 66–77.

## Appendix D

### Sensitivity to the choice of the spatial weight matrix

Appendix D contains the results of the first robustness check, estimating the model in Eq.(3)-(4) using different specifications for the spatial weight matrix  $W$ . Specifically,  $W_{200t}$ ,  $W_{100t}$ ,  $W_{50t}$  and  $W_{30t}$  indicate inverse distance spatial weight matrices truncated at 200, 100, 50 and 30 kilometres respectively, while  $W_{400n}$ ,  $W_{250n}$ ,  $W_{100n}$ ,  $W_{50n}$  and  $W_{30n}$  stand for inverse distance spatial weight matrices considering only the 400, 250, 100, 50 and 30 nearest neighbors, respectively. Table D1 shows some descriptive statistics on the different spatial weight matrix used in this analysis. The most relevant difference between the two types of spatial weight matrices previously defined concern that fact that the number of neighbors changes for each spatial units using a truncated  $W$  while every unit has the same number of neighbors in the second case. Moreover, starting from a truncation point of 50 km some units are considered as islands (i.e. observations with no neighbors). However, in both cases, the mean spatial weight tends to decrease as the radius or the number of nearest neighbors considered decreases. Finally, Table D2 and Table D3 show the estimation results of the SDF-STE model for each spatial weight matrix defined before. Specifically, Table D2 presents the results using different inverse distance truncated  $W$  while Table D3 shows the estimation results for inverse distance  $W$  considering only the  $n$  nearest neighbors.



**Table D1:** Spatial Weight Matrices: Descriptive Statistics

	Mean $w_{ij}$	Mean $d_{ij}$	Min neigh.	10th perc.	Mean neigh.	90th perc.	Max neigh.	Islands
<b>W</b>	0.0093	107.52	5409	5409	5409	5409	5409	0
<b>W200t</b>	0.0074	20.58	14	247	986.03	1603	2030	0
<b>W100t</b>	0.0066	9.55	5	104	341.29	526	733	0
<b>W50t</b>	0.0062	4.96	0	34	158.61	332	442	1
<b>W30t</b>	0.0059	3.11	0	12	98.53	263	412	4
<b>W400n</b>	0.0079	10.35	400	400	400	400	400	0
<b>W250n</b>	0.0073	7.21	250	250	250	250	250	0
<b>W100n</b>	0.0059	3.66	100	100	100	100	100	0
<b>W50n</b>	0.0049	2.19	50	50	50	50	50	0
<b>W30n</b>	0.0042	1.49	30	30	30	30	30	0

Mean  $w_{ij}$  is calculated before row-normalization;  
neigh=neighbors; perc.=percentile;  
Mean  $d_{ij}$  between neighbours expressed in km.

**Table D2: Sensitivity to the Choice of W: Inverse Distance Truncated W**

	<b>W</b>		<b>W200t</b>		<b>W100t</b>		<b>W50t</b>		<b>W30t</b>	
	<i>Coeff.</i>	<i>t-stats</i>	<i>Coeff.</i>	<i>t-stats</i>	<i>Coeff.</i>	<i>t-stats</i>	<i>Coeff.</i>	<i>t-stats</i>	<i>Coeff.</i>	<i>t-stats</i>
$\beta_0$	5.467***	13.71	5.876	14.86***	6.058***	15.07	6.244***	15.27	6.439***	15.35
$\beta_L$	0.662***	84.87	0.661***	84.76	0.662***	84.83	0.661***	84.76	0.661***	84.71
$\beta_K$	0.077***	17.20	0.078***	17.29	0.078***	17.27	0.078***	17.29	0.078***	17.29
$\beta_{LL}$	0.057***	40.57	0.057***	40.71	0.057***	40.64	0.057***	40.64	0.057***	40.57
$\beta_{KK}$	0.010***	22.50	0.010***	25.00	0.01***	25.00	0.01***	25.00	0.01***	25.00
$\beta_{LK}$	-0.029***	-29.00	-0.030***	-29.60	-0.030***	-29.70	-0.030***	-29.60	-0.029***	-29.50
$\beta_t$	-0.005	-1.09	-0.001	-0.24	-0.001	-0.20	-0.001	-0.09	0.001	0.23
$\beta_{2t}$	-0.001*	-1.25	-0.001**	-2.00	-0.001**	-2.00	-0.001**	-2.00	-0.001**	-2.00
$\beta_{tL}$	0.004***	5.00	0.004***	5.00	0.004***	5.00	0.004***	5.00	0.004***	5.00
$\beta_{tK}$	0.001***	1.25	0.001***	3.00	0.001***	3.00	0.001***	3.00	0.001***	3.00
$\rho$	0.351***	18.30	0.253***	18.06	0.212***	17.82	0.171***	9.23	0.132***	15.13
$\theta_L$	-0.216***	-11.33	-0.146***	-10.57	-0.119***	-10.15	-0.091***	-9.23	-0.064***	-7.47
$\theta_K$	0.006	0.78	0.008*	1.42	0.007*	1.49	0.005*	1.26	0.001	0.14
$\phi_0$	5.251***	13.23	5.402***	13.71	5.493***	13.70	5.575***	13.67	5.625***	13.43
$\phi_{hum}$	-0.647***	-208.58	-0.646***	-208.26	-0.645***	-208.06	-0.644***	-207.74	-0.644***	-207.87
$\phi_{Int}$	-0.115***	-29.54	-0.116***	-29.62	-0.115***	-29.59	-0.115***	-29.56	-0.115***	-29.56
$\phi_{pat}$	-0.054***	-9.53	-0.054***	-9.51	-0.053***	-9.33	-0.053***	-9.30	-0.053***	-9.30
$\phi_{trad}$	-0.038***	-5.78	-0.038***	-5.89	-0.039***	-5.92	-0.039***	-6.00	-0.038***	-5.91
$\phi_{size}$	-0.037***	-10.57	-0.036***	-10.37	-0.037***	-10.66	-0.038***	-10.94	-0.039***	-11.09
$\phi_{dsize}$	0.081***	5.66	0.080***	5.58	0.079***	5.51	0.080***	5.54	0.077***	5.35
$\delta_{hum}$	0.160***	7.31	0.116***	7.55	0.085***	6.50	0.055***	4.95	0.036***	3.74
$\delta_{Int}$	-0.072***	-2.71	-0.061***	-3.23	-0.060***	-3.71	-0.052***	-3.84	-0.032***	-2.69
$\delta_{pat}$	0.099***	2.80	0.094***	3.51	0.070***	3.03	0.037**	1.90	0.022*	1.31
$\delta_{trad}$	0.039	0.97	0.033	1.10	0.035*	1.34	0.020	0.89	0.009	0.44
$\delta_{size}$	-0.013	-0.70	-0.026**	-2.02	-0.020**	-1.73	-0.012	-1.20	-0.015**	-1.77
$\delta_{dsize}$	-0.358***	-3.85	-0.230***	-3.43	-0.182***	-3.14	-0.134***	-2.73	-0.108***	-2.59
$\phi_{city}$	-0.064***	-5.94	-0.069***	-6.46	-0.066***	-6.22	-0.062***	-5.87	-0.059***	-5.63
$\phi_{cult}$	0.016*	1.43	0.019**	1.62	0.017*	1.52	0.019**	1.63	0.024**	2.13
$\phi_{sea}$	-0.095***	-8.89	-0.096***	-8.94	-0.095***	-8.83	-0.089***	-8.41	-0.083***	-7.81
$\phi_{lake}$	-0.139***	-9.06	-0.133***	-8.67	-0.131***	-8.55	-0.134***	-8.76	-0.133***	-8.67
$\phi_{mou}$	-0.007	-0.45	-0.004	-0.25	-0.011	-0.72	-0.019	-1.25	-0.012	-0.78
$\phi_{csea}$	-0.085***	-8.12	-0.087***	-8.38	-0.086***	-8.30	-0.080***	-7.71	-0.072***	-6.96
$\phi_{cmou}$	-0.065***	-5.04	-0.055***	-4.32	-0.060***	-4.66	-0.063***	-4.84	-0.061***	-4.77
$\phi_{more}$	-0.047***	-3.77	-0.045***	-3.67	-0.045***	-3.68	-0.044***	-3.59	-0.040***	-3.24
$\phi_{notcat}$	0.052***	4.71	0.053***	4.78	0.053***	4.84	0.053***	4.86	0.0542***	4.97
$\phi_{notur}$	0.015	0.48	0.012	0.39	0.013	0.41	0.012	0.38	0.011	0.34
$\phi_{therm}$	omit.		omit.		omit.		omit.		omit.	
$\sigma^2$	0.199	-	0.199	-	0.199	-	0.199	-	0.200	-
$\lambda$	0.879	-	0.888	-	0.886	-	0.885	-	0.882	-

\*\*\*: *pvalue* ≤ 0.01; \*\*: *pvalue* ≤ 0.05; \*: *pvalue* ≤ 0.10; omit=omitted

**Table D3: Sensitivity to the Choice of W: Nearest Neighbors**

	W400n		W250n		W100n		W50n		W30n	
	Coeff.	t-stats	Coeff.	t-stats	Coeff.	t-stats	Coeff.	t-stats	Coeff.	t-stats
$\beta_0$	5.969**	15.63	6.046***	16.27	6.205***	16.31	6.359***	15.99	6.429***	16.38
$\beta_L$	0.662***	84.87	0.662***	84.87	0.662***	84.87	0.662***	84.87	0.663***	85.00
$\beta_K$	0.078***	17.33	0.078***	17.33	0.078***	17.33	0.078***	17.33	0.078***	17.33
$\beta_{LL}$	0.057***	40.71	0.057***	40.71	0.057***	40.71	0.057***	40.71	0.057***	40.71
$\beta_{KK}$	0.010***	25.00	0.010***	25.00	0.010***	25.00	0.010***	25.00	0.010***	25.00
$\beta_{LK}$	-0.030***	-30.00	-0.03***	-30.00	-0.03***	-30.00	-0.03***	-30.00	-0.03***	-30.00
$\beta_t$	-0.001**	-0.23	-0.001	-0.23	-0.001	-0.23	-0.001	-0.23	-0.001	-0.23
$\beta_{2t}$	-0.001***	-2.50	-0.001***	-2.50	-0.001***	-2.50	-0.001***	-2.50	-0.001***	-2.50
$\beta_{tL}$	0.004***	5.00	0.004***	5.00	0.004***	5.00	0.004***	5.00	0.004***	5.00
$\beta_{tK}$	0.001***	2.50	0.001***	2.50	0.001***	2.50	0.001***	2.50	0.001***	2.50
$\rho$	0.234***	18.14	0.217***	18.08	0.180***	17.31	0.153***	19.15	0.131***	15.23
$\theta_L$	-0.132***	-10.39	-0.122***	-10.34	-0.101***	-9.90	-0.081***	-10.98	-0.067***	-7.98
$\theta_K$	0.005***	0.94	0.004	0.82	0.004	0.93	0.003	1.03	0.006**	1.71
$\phi_0$	5.442**	14.29	5.481***	14.80	5.546***	14.62	5.633***	14.33	5.661***	14.45
$\phi_{hum}$	-0.645***	-208.06	-0.645***	-208.06	-0.644***	-207.74	-0.644***	-207.74	-0.644***	-207.74
$\phi_{Int}$	-0.115***	-29.49	-0.115***	-29.49	-0.115***	-29.49	-0.116***	-29.49	-0.116***	-29.74
$\phi_{pat}$	-0.053***	-9.30	-0.053***	-9.30	-0.053***	-9.30	-0.052***	-9.30	-0.052***	-9.12
$\phi_{trad}$	-0.038***	-5.85	-0.038***	-5.85	-0.039***	-6.00	-0.039***	-6.00	-0.039***	-6.00
$\phi_{size}$	-0.037***	-10.57	-0.038***	-10.86	-0.038***	-10.86	-0.038***	-10.86	-0.038***	-10.86
$\phi_{dsize}$	0.080***	5.56	0.080***	5.56	0.079***	5.49	0.079***	5.49	0.078***	5.42
$\delta_{hum}$	0.097***	6.78	0.084***	6.27	0.062***	5.30	0.044***	5.90	0.031***	3.23
$\delta_{Int}$	-0.050***	-2.91	-0.052***	-3.25	-0.057***	-4.04	-0.044***	-4.56	-0.039***	-3.45
$\delta_{pat}$	0.067***	2.73	0.046**	2.02	0.024	1.22	0.020*	1.37	0.012	0.75
$\delta_{trad}$	0.045**	1.58	0.044**	1.65	0.028	1.20	0.033*	1.35	0.032**	1.70
$\delta_{size}$	-0.017***	-1.39	-0.017*	-1.48	-0.017**	-1.65	-0.021**	-1.81	-0.017**	-1.98
$\delta_{dsize}$	-0.196**	-3.06	-0.181***	-3.01	-0.151***	-2.87	-0.155***	-3.23	-0.144***	-3.42
$\phi_{city}$	-0.067***	-6.32	-0.064***	-6.04	-0.058***	-5.52	-0.061***	-5.52	-0.063***	-6.06
$\phi_{cult}$	0.015***	1.32	0.014	1.23	0.017*	1.49	0.017*	1.49	0.018*	1.58
$\phi_{sea}$	-0.096***	-8.97	-0.095***	-8.88	-0.09***	-8.49	-0.089***	-8.49	-0.088***	-8.30
$\phi_{lake}$	-0.129***	-8.43	-0.133***	-8.69	-0.135***	-8.82	-0.135***	-8.82	-0.137***	-8.95
$\phi_{mou}$	-0.008**	-0.51	-0.015	-0.95	-0.017	-1.08	-0.018	-1.08	-0.016	-1.01
$\phi_{csea}$	-0.088***	-8.46	-0.087***	-8.37	-0.081***	-7.79	-0.079***	-7.86	-0.077***	-7.48
$\phi_{cmou}$	-0.060***	-4.65	-0.061***	-4.73	-0.063***	-4.88	-0.064***	-4.92	-0.063***	-4.92
$\phi_{more}$	-0.045***	-3.63	-0.045***	-3.63	-0.044***	-3.58	-0.044***	-3.58	-0.044***	-3.58
$\phi_{notcat}$	0.053***	4.82	0.053***	4.82	0.050***	4.55	0.046***	4.55	0.046***	4.18
$\phi_{notur}$	0.012*	0.38	0.006	0.19	0.002	0.06	-0.002	-0.06	-0.002	-0.06
$\phi_{therm}$	omit.		omit.		omit.		omit.		omit.	
$\sigma^2$	0.1994		0.199		0.199		0.199		0.199	
$\lambda$	0.8881		0.895		0.896		0.916		0.908	

\*\*\* :  $pvalue \leq 0.01$ ; \*\* :  $pvalue \leq 0.05$ ; \* :  $pvalue \leq 0.10$ ; omit=omitted

## Appendix E

### Robustness Check

Appendix E contains the results of different robustness checks. First, we check whether our non-spatial estimates are affected by the presence of endogeneity or by unmodelled individual heterogeneity. Specifically, Table E1 shows our non-spatial estimates (SF-TE) corresponding to the Battese and Coelli (1995) specification compared to the estimates obtained using the fixed effects model by Greene (2005a). Moreover, Table E2 shows the SDF-STE model estimates introducing the one year and two year lagged input variables and lagged determinants of inefficiency aiming to partially control for possible endogeneity issues. In Table E3 we both control for individual fixed effects and for possible simultaneity issues using both a fixed effects model and lagged variables. The results show that our baseline results are robust to the different modelling approaches.

In addition, we perform a further robustness check in which we estimate our SDF-STE model including other specific territorial indicators at municipal level to capture location specific attributes. In particular, we include in our model specification three variables: *5Stars* that is the ratio between five stars hotels in the municipality and the overall number of accommodation facilities aiming at capturing for high-quality destinations; *Income* that is the ratio between the income of each municipality and the population that is an indicator of wealth; and *Empl* that is the ratio between the number of employees in the active firms of the municipality and the number of active firms proxying the presence of infrastructures in the local territory. The results shown in Table E4 indicate that none of the three territorial indicators considered results to be significantly different from zero at a 1% significance level. Moreover, our baseline estimates are robust to the inclusion of other location specific variables.

Finally, we test the robustness of our estimates to a different indicator for size, that is the logarithm of total assets. In this case, *dSize* is excluded from the specification because we do not need to control for the zero values in *Size* since total assets only take positive values. The results shown in Table E5 indicate that the differences are negligible and thus, the logarithm of the number of managers provides a robust measure of firm size.

- Battese, G.E. and T.J. Coelli (1995). "A model for technical inefficiency effects in a stochastic frontier production function for panel data". *Empirical Economics* 20(2), 325-332.
- Greene, W. (2005a). "Fixed and Random Effects in Stochastic Frontier Models". *Journal of Productivity Analysis* 23, 7-32.

**Table E1.** Comparison between SF-TE model and non-spatial SF model with fixed effects

	SF-TE		Fixed Effects	
	<i>Coeff.</i>	<i>t-stats</i>	<i>Coeff.</i>	<i>t-stats</i>
$\beta_0$	6.665***	7.68	6.826***	5.17
$\beta_L$	0.665***	85.11	0.652***	84.51
$\beta_K$	0.077***	17.04	0.081***	18.17
$\beta_{LL}$	0.057***	40.43	0.054***	39.34
$\beta_{KK}$	0.010***	25.00	0.010***	24.07
$\beta_{LK}$	-0.029***	-26.46	-0.028***	-26.42
$\beta_t$	0.005	0.96	0.006	1.48
$\beta_{2t}$	-0.001**	-2.75	-0.001***	-3.35
$\beta_{tL}$	0.004***	4.88	0.004***	5.28
$\beta_{tK}$	0.001***	3.50	0.001***	3.55
$\phi_0$	5.363***	6.22	5.206***	6.13
$\phi_{hum}$	-0.653***	-210.74	-0.633***	-205.52
$\phi_{Int}$	-0.116***	-29.05	-0.110***	-27.61
$\phi_{pat}$	-0.054***	-9.38	-0.045***	-7.47
$\phi_{trad}$	-0.039***	-5.91	-0.044***	-6.60
$\phi_{size}$	-0.040***	-11.54	-0.048***	-13.03
$\phi_{dsize}$	0.069***	4.73	0.067***	4.60
Destination dummies	yes	-	yes	-
$\sigma^2$	0.20	-	0.16	-
$\lambda$	0.86	-	0.85	-

\*\*\* : *pvalue*  $\leq$  0.01; \*\* : *pvalue*  $\leq$  0.05; \* : *pvalue*  $\leq$  0.10

**Table E2.** Comparison between SDF-STE model and SDF-STE model with lagged variables

	SDF-STE			Lag-1			Lag-2	
	Coeff.	t-stats		Coeff.	t-stats		Coeff.	t-stats
$\beta_0$	5.467***	13.71	$\beta_0$	5.947***	15.62	$\beta_0$	5.372***	12.57
$\beta_L$	0.662***	84.87	$\beta_L(t-1)$	0.660***	70.26	$\beta_L(t-2)$	0.635***	53.80
$\beta_K$	0.077***	17.20	$\beta_K(t-1)$	0.058***	10.87	$\beta_K(t-2)$	0.053***	8.22
$\beta_{LL}$	0.057***	40.57	$\beta_{LL}(t-1)$	0.057***	35.38	$\beta_{LL}(t-2)$	0.059***	32.72
$\beta_{KK}$	0.010***	22.50	$\beta_{KK}(t-1)$	0.011***	27.50	$\beta_{KK}(t-2)$	0.011***	22.40
$\beta_{LK}$	-0.029***	-29.00	$\beta_{LK}(t-1)$	-0.031***	-25.92	$\beta_{LK}(t-2)$	-0.032***	-24.46
$\beta_t$	0.005	1.09	$\beta_t$	0.009	1.27	$\beta_t$	0.043***	4.00
$\beta_{2t}$	-0.001**	-1.25	$\beta_{2t}$	-0.003***	-5.00	$\beta_{2t}$	-0.006***	-7.13
$\beta_{tL}$	0.004***	5.00	$\beta_{tL}(t-1)$	0.005***	4.80	$\beta_{tL}(t-2)$	0.006***	4.69
$\beta_{tK}$	0.001***	3.00	$\beta_{tK}(t-1)$	0.004***	7.20	$\beta_{tK}(t-2)$	0.004***	6.50
$\rho$	0.351***	18.30	$\rho$	0.359***	18.03	$\rho$	0.367***	17.40
$\theta_L$	-0.216***	-11.33	$\theta_L(t-1)$	-0.219***	-10.77	$\theta_L(t-2)$	-0.219***	-10.03
$\theta_K$	0.006	0.78	$\theta_K(t-1)$	0.007	0.78	$\theta_K(t-2)$	0.006	0.62
$\phi_0$	5.251***	13.23	$\phi_0$	5.749***	15.44	$\phi_0$	5.269***	12.52
$\phi_{hum}$	-0.647***	-208.58	$\phi_{hum}(t-1)$	-0.602***	-176.94	$\phi_{hum}(t-2)$	-0.583***	-149.36
$\phi_{Int}$	-0.115***	-29.54	$\phi_{Int}(t-1)$	-0.126***	-27.98	$\phi_{Int}(t-2)$	-0.129***	-25.31
$\phi_{pat}$	-0.054***	-9.53	$\phi_{pat}$	-0.081***	-12.42	$\phi_{pat}$	-0.106***	-14.28
$\phi_{trad}$	-0.038***	-5.78	$\phi_{trad}$	-0.057***	-7.70	$\phi_{trad}$	-0.077***	-9.12
$\phi_{size}$	-0.037***	-10.57	$\phi_{size}$	-0.045***	-11.54	$\phi_{size}$	-0.054***	-11.98
$\phi_{dsize}$	0.081***	5.66	$\phi_{dsize}$	0.082***	5.03	$\phi_{dsize}$	0.081***	4.39
$\delta_{hum}$	0.160***	7.31	$\delta_{hum}(t-1)$	0.121***	5.14	$\delta_{hum}(t-2)$	0.108***	4.15
$\delta_{Int}$	-0.072***	-2.71	$\delta_{Int}(t-1)$	-0.051**	-1.72	$\delta_{Int}(t-2)$	-0.030	-0.89
$\delta_{pat}$	0.099***	2.80	$\delta_{pat}$	0.127***	3.20	$\delta_{pat}$	0.146***	3.21
$\delta_{trad}$	0.039	0.97	$\delta_{trad}$	0.050	1.10	$\delta_{trad}$	0.039	0.75
$\delta_{size}$	-0.013	-0.70	$\delta_{size}$	-0.023	-1.10	$\delta_{size}$	-0.032*	-1.34
$\delta_{dsize}$	-0.358***	-3.85	$\delta_{dsize}$	-0.327***	-3.11	$\delta_{dsize}$	-0.307**	-2.55
Destination dummies	yes	-	Destination dummies	yes	-	Destination dummies	yes	-
$\sigma^2$	0.199	-	$\sigma^2$	0.225	-	$\sigma^2$	0.258	-
$\lambda$	0.879	-	$\lambda$	0.371	-	$\lambda$	0.515	-

\*\*\* :  $pvalue \leq 0.01$ ; \*\* :  $pvalue \leq 0.05$ ; \* :  $pvalue \leq 0.10$

**Table E3.** Comparison between SF-TE model and non-spatial SF model with lagged variables and fixed effects

	SF-TE			Fixed Effects and Lag-1			Fixed Effects and Lag-2	
	<i>Coeff.</i>	<i>t-stats</i>		<i>Coeff.</i>	<i>t-stats</i>		<i>Coeff.</i>	<i>t-stats</i>
$\beta_0$	6.665***	7.68	$\beta_0$	6.737***	16.21	$\beta_0$	6.181***	26.16
$\beta_L$	0.665***	85.11	$\beta_L(t-1)$	0.652***	74.5	$\beta_L(t-2)$	0.634***	63.08
$\beta_K$	0.077***	17.04	$\beta_K(t-1)$	0.062***	12.28	$\beta_K(t-2)$	0.059***	10.13
$\beta_{LL}$	0.057***	40.43	$\beta_{LL}(t-1)$	0.054***	35.15	$\beta_{LL}(t-2)$	0.058***	32.73
$\beta_{KK}$	0.010***	25.00	$\beta_{KK}(t-1)$	0.011***	24.6	$\beta_{KK}(t-2)$	0.012***	22.77
$\beta_{LK}$	-0.029***	-26.46	$\beta_{LK}(t-1)$	-0.030***	-25.15	$\beta_{LK}(t-2)$	-0.031***	-23.31
$\beta_t$	0.005	0.96	$\beta_t$	0.044***	7.70	$\beta_t$	0.101***	11.25
$\beta_{2t}$	-0.001**	-2.75	$\beta_{2t}$	-0.005***	-11.64	$\beta_{2t}$	-0.009***	-13.68
$\beta_{tL}$	0.004***	4.88	$\beta_{tL}(t-1)$	0.003***	3.76	$\beta_{tL}(t-2)$	0.003***	2.87
$\beta_{tK}$	0.001***	3.50	$\beta_{tK}(t-1)$	0.004***	8.21	$\beta_{tK}(t-2)$	0.004***	6.97
$\phi_0$	5.363***	6.22	$\phi_0$	4.974***	14.62	$\phi_0$	5.633***	14.33
$\phi_{hum}$	-0.653***	-210.74	$\phi_{hum}(t-1)$	-0.578***	-166.44	$\phi_{hum}(t-2)$	-0.555***	-141.30
$\phi_{Int}$	-0.116***	-29.05	$\phi_{Int}(t-1)$	-0.123***	-27.28	$\phi_{Int}(t-2)$	-0.128***	-24.71
$\phi_{pat}$	-0.054***	-9.38	$\phi_{pat}$	-0.073***	-10.8	$\phi_{pat}$	-0.099***	-12.96
$\phi_{trad}$	-0.039***	-5.91	$\phi_{trad}$	-0.067***	-8.83	$\phi_{trad}$	-0.088***	-10.20
$\phi_{size}$	-0.040***	-11.54	$\phi_{size}$	-0.059***	-14.19	$\phi_{size}$	-0.070***	-14.75
$\phi_{dsize}$	0.069***	4.73	$\phi_{dsize}$	0.068***	4.13	$\phi_{dsize}$	0.066***	3.49
Destination dummies	yes	-	Destination dummies	yes	-	Destination dummies	yes	-
$\sigma^2$	0.20	-	$\sigma^2$	0.18	-	$\sigma^2$	0.21	-
$\lambda$	0.86	-	$\lambda$	0.84	-	$\lambda$	0.85	-

\*\*\* : *pvalue* ≤ 0.01; \*\* : *pvalue* ≤ 0.05; \* : *pvalue* ≤ 0.10



**Table E4.** Including other specific territorial indicators

	SDF-STE		Extended specification	
	<i>Coeff.</i>	<i>t-stats</i>	<i>Coeff.</i>	<i>t-stats</i>
$\beta_0$	5.467***	13.71	5.462***	9.54
$\beta_L$	0.662***	84.87	0.658***	87.75
$\beta_K$	0.077***	17.20	0.078***	19.90
$\beta_{LL}$	0.057***	40.57	0.057***	40.71
$\beta_{KK}$	0.010***	22.50	0.010***	33.00
$\beta_{LK}$	-0.029***	-29.00	-0.029***	-32.56
$\beta_t$	0.005	1.09	-0.003	-0.70
$\beta_{2t}$	-0.001**	-1.25	-0.001*	-1.50
$\beta_{tL}$	0.004***	5.00	0.004***	5.00
$\beta_{tK}$	0.001***	3.00	0.001***	3.00
$\rho$	0.351***	18.30	0.343***	18.22
$\theta_L$	-0.216***	-11.33	-0.231***	-12.16
$\theta_K$	0.006	0.78	0.013**	1.63
$\phi_0$	5.251***	13.23	5.189***	9.17
$\phi_{hum}$	-0.647***	-208.58	-0.644***	-207.87
$\phi_{Int}$	-0.115***	-29.54	-0.114***	-29.10
$\phi_{pat}$	-0.054***	-9.53	-0.054***	-9.42
$\phi_{trad}$	-0.038***	-5.78	-0.038***	-5.89
$\phi_{size}$	-0.037***	-10.57	-0.035***	-10.00
$\phi_{dsize}$	0.081***	5.66	0.084***	5.86
$\delta_{hum}$	0.160***	7.31	0.190***	8.66
$\delta_{Int}$	-0.072***	-2.71	-0.063**	-2.42
$\delta_{pat}$	0.099***	2.80	0.077**	2.17
$\delta_{trad}$	0.039	0.97	0.030	0.73
$\delta_{size}$	-0.013	-0.70	-0.042**	-2.27
$\delta_{dsize}$	-0.358***	-3.85	-0.312***	-3.35
$\phi_{city}$	-0.064***	-5.94	-0.043***	-3.97
$\phi_{cult}$	0.016*	1.43	0.031**	2.72
$\phi_{sea}$	-0.095***	-8.89	-0.103***	-9.66
$\phi_{lake}$	-0.139***	-9.06	-0.129***	-8.46
$\phi_{mou}$	-0.007	-0.45	0.001	0.06
$\phi_{csea}$	-0.085***	-8.12	-0.089***	-8.46
$\phi_{cmou}$	-0.065***	-5.04	-0.037**	-2.88
$\phi_{more}$	-0.047***	-3.77	-0.041***	-3.29
$\phi_{notcat}$	0.052***	4.71	0.054***	4.86
$\phi_{notur}$	0.015	0.48	0.025	0.81
$\phi_{5stars}$	-	-	-0.006	-0.10
$\phi_{inc}$	-	-	-0.000	-0.01
$\phi_{empl}$	-	-	-0.005	-0.25
$\sigma^2$	0.199	-	0.199	-
$\lambda$	0.879	-	0.451	-

\*\*\* :  $pvalue \leq 0.01$ ; \*\* :  $pvalue \leq 0.05$ ; \* :  $pvalue \leq 0.10$

**Table E5.**  
a different  
size

	Firm size as num. of managers		Firm size as total assets	
	Coeff.	t-stats	Coeff.	t-stats
$\beta_0$	5.467***	13.71	6.475***	10.29
$\beta_L$	0.662***	84.87	0.664***	87.37
$\beta_K$	0.077***	17.20	0.068***	15.84
$\beta_{LL}$	0.057***	40.57	0.058***	44.85
$\beta_{KK}$	0.010***	22.50	0.007***	17.50
$\beta_{LK}$	-0.029***	-29.00	-0.031***	-30.50
$\beta_t$	0.005	1.09	-0.012***	-2.73
$\beta_{2t}$	-0.001**	-1.25	-0.001*	-1.50
$\beta_{tL}$	0.004***	5.00	0.003***	3.50
$\beta_{tK}$	0.001***	3.00	0.002***	5.00
$\rho$	0.351***	18.30	0.303***	15.31
$\theta_L$	-0.216***	-11.33	-0.173***	-9.14
$\theta_K$	0.006	0.78	-0.013*	-1.45
$\phi_0$	5.251***	13.23	6.308***	10.04
$\phi_{hum}$	-0.647***	-208.58	-0.642***	-213.97
$\phi_{Int}$	-0.115***	-29.54	-0.107***	-28.03
$\phi_{pat}$	-0.054***	-9.53	-0.046***	-8.20
$\phi_{trad}$	-0.038***	-5.78	-0.042***	-6.59
$\phi_{size}$	-0.037***	-10.57	-0.064***	-53.58
$\phi_{dsize}$	0.081***	5.66	-	-
$\delta_{hum}$	0.160***	7.31	0.110***	5.13
$\delta_{Int}$	-0.072***	-2.71	-0.057**	-2.24
$\delta_{pat}$	0.099***	2.80	0.094***	2.73
$\delta_{trad}$	0.039	0.97	0.066*	1.68
$\delta_{size}$	-0.013	-0.70	-0.025***	-3.28
$\delta_{dsize}$	-0.358***	-3.85	-	-
$\phi_{city}$	-0.064***	-5.94	-0.053***	-5.05
$\phi_{cult}$	0.016*	1.43	0.013	1.13
$\phi_{sea}$	-0.095***	-8.89	-0.075***	-7.17
$\phi_{lake}$	-0.139***	-9.06	-0.119***	-7.99
$\phi_{mou}$	-0.007	-0.45	-0.018	-1.18
$\phi_{csea}$	-0.085***	-8.12	-0.067***	-6.56
$\phi_{cmou}$	-0.065***	-5.04	-0.078**	-6.26
$\phi_{more}$	-0.047***	-3.77	-0.031***	-2.58
$\phi_{notcat}$	0.052***	4.71	0.042***	3.89
$\phi_{notur}$	0.015	0.48	0.043*	1.42
$\sigma^2$	0.199	-	0.189	-
$\lambda$	0.879	-	0.473	-

Robustness to  
definition for

\*\*\* :  $pvalue \leq 0.01$ ; \*\* :  $pvalue \leq 0.05$ ; \* :  $pvalue \leq 0.10$