



ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rsea20

The effect of agglomeration economies and geography on the survival of accommodation businesses in Sicily

Davide Piacentino, Martina Aronica, Diego Giuliani, Andrea Mazzitelli & Maria Francesca Cracolici

To cite this article: Davide Piacentino, Martina Aronica, Diego Giuliani, Andrea Mazzitelli & Maria Francesca Cracolici (2020): The effect of agglomeration economies and geography on the survival of accommodation businesses in Sicily, Spatial Economic Analysis, DOI: 10.1080/17421772.2020.1836389

To link to this article: https://doi.org/10.1080/17421772.2020.1836389



View supplementary material

•	0

Published online: 05 Nov 2020.

ம	

Submit your article to this journal 🖸

Article views: 197



View related articles



🌔 View Crossmark data 🗹





Check for updates

The effect of agglomeration economies and geography on the survival of accommodation businesses in Sicily

Davide Piacentino ^a, Martina Aronica ^b, Diego Giuliani ^c, Andrea Mazzitelli ^d and Maria Francesca Cracolici ^e

ABSTRACT

The study explores the geographical pattern of the accommodation industry in the Italian insular region of Sicily, focusing on the determinants of the risk of market exit. We adopt a standard framework of business survival analysis where agglomeration economies play an important role. We then extend the analysis by considering the role of geography to explore whether the risk of market exit depends on nearness to desirable amenities. The geography is here measured by the distance from the coast and the altitude of the place where the firm is located. When we look at the entire population of accommodation firms that started between 2010 and 2014, we find evidence that the risk of failure increases for those which are over 2 km from the coast.

KEYWORDS

firm survival, tourism industry, geographical location, insular regions

JEL C41, R12, Z30 HISTORY Received 15 July 2019; in revised form 2 October 2020

INTRODUCTION

Tourism makes a major contribution to the economy of many countries and may also be a tool to foster regional development (e.g., Balaguer & Cantavella-Jorda, 2002; Brida et al., 2014; Li et al., 2018; Neves Sequeira & Maçãs Nunes, 2011). Researchers have also presented evidence on the potential of tourism growth for fostering the development of rural and peripheral regions, as well as for reducing disparities in less developed areas within regions (Brau et al., 2007; Liu et al., 2017). In rural and peripheral areas that have little in the way of industry, tourism may be an instrument of development (Brouder & Eriksson, 2013; Liu et al., 2017).

CONTACT

^a avide.piacentino@unipa.it

^b martina.aronica@unipa.it

Supplemental data for this article can be accessed at https://doi.org/10.1080/17421772.2020.1836389

Department of Economics, Business and Statistics, University of Palermo, Palermo, Italy.

Department of Economics, Business and Statistics, University of Palermo, Palermo, Italy.

c 🔁 diego.giuliani@unitn.it

Department of Economics and Management, University of Trento, Trento, Italy.

d 🖾 a.mazzitelli@unimercatorum.it

Universitas Mercatorum, Rome, Italy.

e (Corresponding author) 😡 mariafrancesca.cracolici@unipa.it

Department of Economics, Business and Statistics, University of Palermo, Palermo, Italy.

Tourism growth is fed by the demand that reflects increasing wealth and economic development, but also depends on the ability of the tourism industry itself to offer new products and experiences that meet the increased demand (Alsos et al., 2014; Solvoll et al., 2015). In spite of that, the literature has mainly focused on the analysis of regional tourism competitiveness from a demand-side perspective, while less attention has been paid to the role of entrepreneurs and new business formation as key factors in encouraging competitiveness (e.g., Chhetri et al., 2008; Estevão & Ferreira, 2009; Fritsch & Schroeter, 2011).

For example, an issue that merits investigation is the ability of new accommodation businesses to survive in the market, which is a direct measure of regional tourism competitiveness. Actually, while this aspect has received little attention in the tourism literature, it is a topic that has been widely explored in industrial and regional economics where the relation between agglomeration economies (localization, urbanization and diversification) and firm survival is a common starting point of empirical analyses (e.g., Howell et al., 2018; Renski, 2011).

Despite this interest, no single study in tourism economics to the best of our knowledge has explored the role of agglomeration economies on firm survival. The only attempt could be by Falk (2013), who looks at the influence of the presence of competitors within a distance of ≤ 10 km on the survival probability of ski-lift firms. However, the aim of the study is not to measure localization economies, but congestion effects. To fill this gap in the tourism literature, the aim of our study is twofold. First, it investigates the expected relation between agglomeration economies and the risk of closure of accommodation businesses using some spatial measures at the firm level. In particular, we explore the effects of localization and relatedness by following the approach suggested by Arbia et al. (2015) (see below) - while urbanization is more standardly measured as the population density at municipality level. Second, we extend the analysis by considering the role of geography so as to explore whether the risk of market exit is related to the nearness of desirable amenities. This part of our work is highly original since most of the existing literature totally ignores the crucial role that the spatial location of firms could have in tourism. Indeed, only two studies have moved in this direction (e.g., Falk, 2013; Gémar et al., 2016).

Our study uses Sicilian micro-data, and it considers the distance of accommodation firms from the coast and the altitude of the municipalities where the firms are located to investigate whether their survival is affected by their geographical location. Actually, we would expect this to be the case in an insular region such as Sicily where distance from the coast could well be a key factor of competitiveness.

Summing up, the paper contributes to the literature on tourism economics in several different ways. As mentioned above, it is the first attempt in the literature to look at the relationship between agglomeration economies and firm survival in the tourism industry. Second, there have been few attempts to consider the role of the spatial location of firms on their chances of survival. By using geocoded micro-data our study is able to investigate this very accurately. Finally, we provide novel evidence on the tourism accommodation industry in Sicily, a region which despite being geographically peripheral and viewed as an economic backwater is known internationally as an attractive region offering a range of tourist products (e.g., Cracolici & Nij-kamp, 2008). Additionally, our analysis is particularly accurate since it is not performed on sampling data, but involves the entire population of Sicilian accommodation firms that started in business between 2010 and 2014.

The remainder of the paper is structured as follows. The next section reviews the existing literature. The third section describes the data and presents the methods. The fourth section presents and discusses the empirical findings. Lastly, some final remarks are left for the concluding section.

LITERATURE REVIEW

Firm survival has often been a subject of study in industrial and regional economics and the literature has mainly concentrated on the relation between agglomeration economies and the ability of new firms to survive in the market (e.g., Audretsch & Fritsch, 1994; Fotopoulos & Louri, 2000; Frenken et al., 2015; Howell et al., 2018; Renski, 2011). There is a wide consensus that agglomeration economies have to be measured in a multidimensional way, including at least three different features: localization, urbanization and diversification (Glaeser et al., 1992; Henderson, 2003; Jacobs, 1969; Nakamura, 1985; Rosenthal & Strange, 2001).

Localization economies (or Marshall externalities) derive from the concentration of firms specialized in the same economic activity. Firms may benefit from sharing resources, knowledge and infrastructure characteristic of a given economic activity (e.g., Rosenthal & Strange, 2001, 2004). In the case of the tourism industry, for example, Segarra-Oña et al. (2012) look at the impact of localization economies on the economic performance and competitiveness of Spanish hotels. Peiró-Signes et al. (2015) carry out a similar investigation on US hotels. In the Italian context, Capone and Boix (2008) find that localization economies contribute to the growth of a local tourism production system more than natural resources and endowments,¹ and Lazzeretti and Capone (2009) show that localization externalities in tourism positively influence economic growth.

Urbanization refers to all those economies of scale deriving from operating in a densely populated urban area. Since urbanization economies depend on several factors (e.g., highly skilled workers, local demand, an innovative environment, infrastructures, etc.), the interpretation of their causes is never straightforward. We would generally expect urbanization to exercise a positive influence on firm performance, but this has not been demonstrated in the case of the tourism industry where no previous research looking into the effect of urbanization on firm survival has been carried out.

Diversification (also known as Jacobs-type externalities) refers to the benefit that a firm can obtain if it is located in a place where there is a wide variety of economic activities. The basic idea is that industrial diversity supports the development of new ideas and innovation leading to economic growth (Glaeser et al., 1992; Jacobs, 1969). However, some authors have pointed out the need to disentangle Jacobs-type externalities into two varieties, termed related and unrelated, so as to explain the effect of knowledge spillovers more accurately (e.g., Boschma & Wenting, 2007; Boschma et al., 2012; Delgado et al., 2010; Frenken, Van Oort, & Verburg, 2007; Porter, 2003). The related variety refers to industries that share some complementarities, not just in terms of used inputs and produced outputs, but of competencies and cognitive proximity (Boschma et al., 2012; Boschma & Iammarino, 2009). Learning and knowledge diffusion is more likely to occur between closely related industries (see also Nooteboom, 2000). The idea of related variety is particularly interesting in the case of tourism where the complementarities among industries (e.g., accommodation, beverage and food, cultural activities, etc.) could strongly affect their performance and survival.

Unfortunately, there is no research on the relation between agglomeration economies and firm survival in the tourism industry. We found only two related studies that focus exclusively on single measures of agglomeration economies. Falk (2013) analyses the influence of competitors within a distance of ≤ 10 km on the survival probability of ski-lift firms which could be considered a measure of localization, although the author interprets it as a congestion effect. As far as urbanization is concerned, Gémar et al. (2016) and Falk (2013) include in their analyses some dummy variables by region or macro-area, but this is not the same as measuring the degree of urbanization.

Studies published in the tourism literature have also largely neglected the importance of spatial location on firm survival and this could be particularly relevant in some specific cases such as those of insular and peripheral regions where location-specific characteristics (i.e., amenities) are key factors of competitiveness. Of the few studies along these lines we could mention Gémar et al. (2016), who analyse post-entry performance in the Spanish hotel industry from 1997 to 2009 and find that the risk of failure, more than any other factor, increases if the firm is > 100 km from an airport. Looking at the survival of ski-lift firms in Austria during the period 1995-11, Falk (2013) finds that the risk of failure decreases if the elevation of the ski area is at least 1700 masl. As far as the Italian tourism industry is concerned, only Santarelli (1998) explores the post-entry performance of firms, but looks only indirectly at the role of geography. Exploring the relationship between firm survival performance and the size of new firms in the hospitality industry, the paper finds that the survival pattern is significantly different across Italian regions. This study already highlights the importance of geography in firm survival in the Italian tourism industry, notwithstanding the limitation of assuming that firms are spatially homogeneous within regions. In our analysis of the accommodation industry in Sicily, however, we aim to overcome this limitation by using georeferenced data at firm level and to fill in some of the gaps in the literature.

DATA AND METHODS

Data and variables

The study exploits census data on business demography, collected by the Italian Institute of Statistics (ISTAT). This database is georeferenced and internationally comparable. In particular, our study uses data on Sicilian start-ups in the accommodation industry during the period 2010–14.²

We focus our analysis on accommodation businesses because they are less affected by local demand than tourism-related services such as restaurants and transport, etc. However, we do use data on related industries (restaurants, renting of motor vehicles, libraries, archives, museums, etc.) in order to obtain a measure of relatedness, as we will see below. The data set consists of 660 new accommodation firms in Sicily that started during the period 2010–14. During this period, 146 of these closed within five years.³

In Table 1, we report the list of variables used as covariates in the regression analysis. The first three variables refer to agglomeration economies (localization, relatedness and urbanization). The first two are measured by using the georeferenced microdata, while the third is based on data at the municipality level. *Localization* is measured by a function of the distance⁴ of a given firm *i* in the accommodation industry from all other firms in the same industry (Arbia et al., 2015). High values of *Localization* indicate that the firm *i* is located nearby other firms in the same industry. In the preliminary descriptive analysis, we will group firms into quartiles to illustrate the differences between firms with higher and lower values of *Localization*. In the regression analysis, on the other hand, we use the variable as continuous so as to respect its nature. The expected sign of *Localization* is ambiguous since we can have two opposite effects. A positive effect is expected if spatial spillovers occur due to the presence of a specialized industrial cluster, while a negative effect should be observed if the market suffers from 'congestion' (i. e., too many accommodation businesses in the same area without sufficient demand).

The variable of *Relatedness* is measured by adopting the same approach, but here considering the distance of a firm *i* in the accommodation industry from firms operating in related industries.⁵ We expect relatedness to exert a positive influence on firm survival as suggested by the literature (Nooteboom, 2000). As we did before, we group firms in quartiles only for descriptive aims and use it as continuous in the regression analysis.

Urbanization is defined by adopting the approach suggested by EUROSTAT, i.e., the Degree of Urbanization Classification (DEGURBA). In our case, we consider two categories:

Variable	Description
Localization	Based on the inverse of the spatial distance of a given firm in the accommodation industry
	from all other firms in the same industry. See Appendix A in the supplemental data online
	for details
Relatedness	Based on the inverse of the spatial distance of a given firm in the accommodation industry
	from all other firms in related industries. See Appendix A online for details
Urbanization	Two categories defined by Eurostat's criteria of geographical contiguity and minimum
	population threshold applied to 1 km ² population grid cells (degree of urbanization
	classification – DEGURBA):
	Low level (reference category): rural and intermediate areas
	High level: urban areas
Sales	Annual turnover of the first year of activity grouped into quartiles:
	• Micro (reference category): first quartile (€0–13,107)
	• Small: second quartile (€13,108–28,907)
	• Medium: third quartile (€28,908–75,875)
	• Large: fourth quartile (€75,876–2,262,482)
Distance from the	Three distance categories:
coast	 < 1 km (reference category)
	• 1–2 km
	• > 2 km
Altitude	Three categories:
	• < 300 masl (reference category)
	• 300–500 masl
	• > 500 masl
Accommodation	Two categories: Hotels = 1; other types otherwise

low urbanization if the firm is to be found in a municipality with intermediate or low population
density (i.e., rural and intermediate areas), and high urbanization if the firm is in a densely popu-
lated municipality (i.e., urban areas). Urbanization usually has a positive impact on new business
formation and survival. For example, using Italian data, Piacentino et al. (2017) find that urban-
ization positively influences new business formation not only directly but also indirectly due to its
relation with the innovative activity of firms. However, this result could not be confirmed in our
study. We are exploring accommodation businesses in an insular region where, due to the impor-
tance of seaside tourism, the location choice of a new firm could be dependent more on its dis-
tance from the coast than on that from densely populated areas.

We use its annual turnover after the first year of activity in order to measure firm size (*Size*).⁶ We prefer to use this measure instead of those based on the number of employees, since many seasonal workers in the accommodation industry may not be recorded in our data. In this case, we group firms into quartiles and use the set of dummy variables in the regression analysis to explore the effects of different sizes (i.e., micro, small, medium and large) on survival probability.

As far as the geography is concerned, we measure the distance of each firm from the coast⁷ and the altitude of the municipality where the firm is located. Regarding distance from the coast, we group firms into three categories: < 1, 1-2 and > 2 km.⁸ Similarly, we group firms according to altitude as follows: < 300, 300–500 and > 500 masl. Finally, we include a variable to distinguish hotels from the other types of accommodation.

Appendix B in the supplemental data online reports some descriptive statistics on the variables described above.

Survival analysis

The empirical analysis of factors affecting firm's survival time, that is, the elapsed time between the entry and exit of a firm, is known as *survival* (or *duration*) analysis. This analysis is based on a regression modelling framework that enables us to deal with the inevitable occurrence of censoring, that is, the presence of truncated observations due to the fact that the follow-up time is shorter than that necessary for a firm's exit to be observed.

In practice, survival regression models specify the relationship between explanatory variables and the so-called hazard function $\lambda_i(t)$, which in the empirical circumstance under consideration represents the instantaneous exit rate for firm *i* surviving to time *t*. Therefore, $\lambda_i(t)dt$ provides the probability of firm *i* to close at time *t* given its survival until time *t*. The parametric survival regression models imply specifying a plausible probability distribution of survival times and hence the corresponding functional form of $\lambda_i(t)$. This limitation is overcome in the semi-parametric model proposed by Cox (1972) where no distributional assumptions are formulated. In the Cox model, the hazard function for firm *i* depends on time *t* and a set of *k* explanatory variables (x_1, x_2, \ldots, x_k) as follows:

$$\lambda_i(t) = \lambda_0(t) \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}) \tag{1}$$

where $\lambda_0(t)$ is the baseline hazard; and β_1 , β_2 , ..., β_k denotes, as usual, unknown parameters to be estimated. The functional form of $\lambda_0(t)$ does not need to be specified because it is assumed to be the same for all firms. Indeed, if we consider two generic firms, say firm *i* and firm *l*, the ratio between their hazards is given by:

$$\frac{\lambda_i(t)}{\lambda_l(t)} = \frac{\lambda_0(t)\operatorname{expexp}\left(\beta_1 x_{1i} + \dots + \beta_k x_{ki}\right)}{\lambda_0(t)\operatorname{expexp}\left(\beta_1 x_{1l} + \dots + \beta_k x_{kl}\right)} = \exp[\beta_1(x_{1i} - x_{1l}) + \dots + \beta_k(x_{ki} - x_{kl})]$$
(2)

and hence it is independent of both time and $\lambda_0(\cdot)$. Since the model specified in equation (1) is distribution-free but requires the hazards to be proportional, it is termed the *Cox proportional-hazards model*. The robustness of the proportional hazards assumption can be validated through the Grambsch–Therneau test (Grambsch & Therneau, 1994), as we will see in the next section.

Estimation and inference about the parameters of the Cox model can be performed using the partial likelihood technique (Cox, 1975). However, in the specific empirical context of firm survival, its use may not be proper. This is because the Cox model treats time as measured on the continuous scale, whereas the duration of firms is here observed in discrete units of a year. Such a circumstance leads to a non-negligible number of tied survival times, which makes the partial likelihood construction inappropriate. The problem can be overcome using the approximation of the exact marginal likelihood proposed by Efron (1977).

To take into consideration the fact that firms located in the same area, such as a municipality, may be more alike in their survival times than firms from different areas, the Cox model also allows each firm in a given area to be assigned a common factor (i.e., a random effect) which in this literature is called a 'shared frailty':

$$\lambda_{ij}(t) = \lambda_0(t)\omega_j \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki})$$
(3)

where $\lambda_{ij}(\cdot)$ is the hazard function for the *i*th firm in the *j*th area; and ω_j represents the random effect (or shared frailty) for all firms in the *j*th area.

The ω_j vary from area *j* to another, and the assumed distribution that governs this variability is a gamma distribution with variance, θ , estimated from the data. The inclusion of this area-level random effect accommodates dependence among the survival times of different firms located in the same area and at the same time controls for the presence of unobserved heterogeneity after conditioning on the covariates included in the model.

When geocoded data are observed, as in our case, spatial variation in survival times may be at play; in other words, some spatial clustering of high or low survival times might be relevant. Therefore, the assumption of the independence of survival times may be violated, and residuals may be spatially correlated. However, potential spatial autocorrelation should be mitigated by including in the specification area-level random effects and some covariates of spatial nature such as distance from the coast as well as localization or relatedness. To be sure that our estimates are not affected by the aforementioned problems, we will use the variogram in the next section as a diagnostic tool (Diggle & Ribeiro, 2007).

EMPIRICAL RESULTS

Preliminary analysis

Figure 1 shows the spatial distribution of Sicilian firms in the accommodation industry that started during the period 2010–14. Crosses indicate firms that closed during the period 2011–15,⁹ while circles represent firms that survive in the same period. The size of symbols is



Figure 1. Firm survival in accommodation industry in Sicily during the period 2010–14. Note: Crosses (circles) represent firms that fail (survive). Their size is proportional to the survival time.

proportional to the survival time. We observe a non-homogeneous spatial distribution of firms in terms of both births (i.e., firms entering the market) and deaths (i.e., firms closing). Generally, firms are mainly concentrated along the coast and near the biggest cities; we observe a significant concentration of firms in the north-west and east of Sicily. We observe that firms are mainly located close to coastal and cultural tourist destinations such as Palermo, Trapani and Syracuse, and business cities such as Catania. In the biggest cities, Palermo and Catania, for example, we note a high concentration of both births and deaths. There are also some areas along the coast in the north and south-west that are far from the biggest cities where we observe a high concentration of births and a low concentration of deaths.

As would be expected, there is a high survival rate of accommodation businesses located near the town of Taormina, which is an internationally renowned tourist destination. Finally, we note a significant concentration of deaths in inland areas near the city of Enna. This first visual inspection gives us insights that post-entry firm performance differs across Sicilian areas, while also showing some positive spatial patterns along the coast within a certain distance from the most urbanized cities.

By using the non-parametric approach of Kaplan and Maier (1985), a preliminary survival analysis on the population of Sicilian accommodation firms (i.e., 660 firms) for the period 2010–14 was performed (Table 2). We observe that the probability that an accommodation firm will survive > 1 year is 90%, decreasing to 66% after five years.

Figure 2 shows the survival curve obtained from the Kaplan and Meier approach. In short, we plot the survival probabilities against the years of entry and in the case of more than one curve we use the log-rank test to assess whether these curves are statistically different (Peto et al., 1977). We observe that the probability of survival decreases over the analysed period; in the first two years from entry the risk of closure increases by about 15%. However, the shape of the survival curves differs depending on the characteristics of the firm (Figure 3).¹⁰

Looking at Figure 3(a), we observe differences in the survival probabilities over time across firms belonging to different quartiles of *Localization*. The log-rank test indicates that these curves differ significantly in statistical terms. In particular, we observe that already after the first year of activity the gap between firms in the fourth quartile and the rest of the firms is large and then dramatically increases over time. Specifically, firms with the highest values of *Localization* (fourth quartile) have no likelihood of surviving past the fourth year of activity (see Table C1 in Appendix C in the supplemental data online). This means that firms in the accommodation industry suffer considerably from a sort of 'congestion effect', that is, too much supply for a limited demand in a given spatial area. This result may well be because we are considering the specific case of an insular region where space may matter more than in other places.

Figure 3(b) shows the survival curves built on the *Relatedness* variable. We note that firms in the fourth quartile consistently have the highest survival probability, ranging from 92% after the

Years from entry	Firms at risk	Firm exits	Survival probability	95% lower Cl	95% upper Cl
1	660	63	0.905	0.882	0.927
2	461	46	0.814	0.783	0.847
3	301	21	0.757	0.721	0.796
4	186	8	0.725	0.684	0.769
5	102	8	0.668	0.616	0.725

Table 2. Survival probabilities of Sicilian firmsborn during the period 2010–14 in the accommodation industry.

Note: CI, confidence interval.



Figure 2. Survival curves of Sicilian firms born in the period 2010–14 in the accommodation industry.

first year to 73% after five years. Additionally, we observe that the probability of survival in the first quartile reduces dramatically after the third year in comparison with all the other groups. However, no significant differences across the curves emerge from the log-rank test. Figure 3 (c) compares firms in urban areas with those in rural and intermediate areas. *Urbanization* seems to have a positive impact on the probability of survival, although the log-rank test results indicate that it is not statistically significant. The gap between the two groups becomes larger only after the fourth year.

There are, however, significant differences in the survival curves when we look at the *Sales* variable (Figure 3d). Firms in the third and fourth quartiles (medium and large firms) have significantly higher probabilities of survival than those in the other two quartiles (micro and small firms), and this difference increases over time. After five years from birth, firms in the first quartile are 20% less likely to survive compared with firms in the last quartile (see Table C4 in Appendix C in the supplemental data online).

Figure 3(e, f) shows the curves for the two variables that describe the geographical location of the firms' distance from the coast and altitude. In both cases, we do not find significant differences across groups. However, we observe that after the first year of activity, firms that are < 1 km from the coast have a higher probability of survival. After the second year, the better location seems to be that between 1 and 2 km from the coast. In the case of altitude, it seems that it is better to stay at < 500 masl to increase survival probability. Finally, we compare firms by type of accommodation.

Figure 3(g) notes that hotels always have higher survival probabilities than other types of accommodation, especially from the third year onwards. However, also in this case the differences are not significant.

In this stage of the analysis, we have compared the survival probabilities of firms grouped together by certain characteristics. However, the results obtained here are only preliminary because they do not account for the dependence among the different characteristics. The next section provides a more rigorous analysis by estimating a Cox proportional-hazard model.



Logrank test $X^2 = 2.4$, *p*-value= 0.1

Figure 3. Survival curves by firm characteristics in the accommodation industry in Sicily during the period 2010–14.

Regression analysis

Table 3 reports the estimates of the Cox proportional hazard model.¹¹ In the first column, we provide estimates of a baseline model where only variables related to the agglomeration economies (localization, relatedness and urbanization) are considered (see model 1). In addition, in model 2 we consider the interaction effect of relatedness and localization to explore the relationship between these two measures of agglomeration.

We note that *Localization* has a negative effect since the risk of closure significantly increases for higher values of localization. Although this is not the expected result, it confirms what was obtained in the preliminary analysis. New firms in the accommodation industry seem to suffer when there is a high spatial concentration of firms in the same industry. In the previous section we referred to this result as a 'congestion effect',¹² that is, the demand is not sufficient to satisfy the entirety supply of all new firms. However, there are other potential interpretations of this effect that may be suitable in the case of lagging regions. For example, we can consider the case of an attractive destination that is not well connected by transportation to neighbouring areas where potentially new accommodation businesses could be localized. In other terms, the obstacle to expanding 'borders' of a tourism destination could not be depending on the potential demand but on the support of regional or local infrastructures. This interpretation gives interesting insights for further research. Looking at the interaction effect we also observe that the negative effect on survival of *Localization* is not moderated at higher levels of *Relatedness*.

	Model 1	Model 2	Model 3
Localization	2.01 (0.12)***	0.67 (0.28)**	0.64 (0.29)**
Relatedness	-0.10 (0.08)	-1.86 (0.37)***	-1.92 (0.37)***
Localization \times Relatedness		0.33 (0.07)***	0.34 (0.07)***
Urbanization: low level	Reference	Reference	Reference
Urbanization: high level	0.07 (0.21)	0.07 (0.21)	0.07 (0.24)
Micro			Reference
Small			0.16 (0.23)
Medium			-0.49 (0.27)*
Large			-0.47 (0.28)*
Distance from coast: < 1 km			Reference
Distance from coast: 1–2 km			0.39 (0.28)
Distance from coast: $> 2 \text{ km}$			0.40 (0.22)*
Altitude: < 300 masl			Reference
Altitude: 300–500 masl			-0.25 (0.39)
Altitude: > 500 masl			-0.25 (0.26)
Accommodation: other types			Reference
Accommodation: hotel			0.13 (0.23)
<i>R</i> ²	0.47	0.48	0.50
Maximum R ²	0.93	0.93	0.93
Shared frailty (θ)	0.01**	0.01**	0.01**
PH test (χ^2)	23.95***	14.79***	17.30
Observations	660	660	660

Table 3. Results from the Cox proportional hazard model.

Note: ***p < 0.01, **p < 0.05, *p < 0.1.

As far as *Relatedness* is concerned, we find a positive effect on firm survival, that is, the high spatial concentration of related industries reduces the risk of failure. More precisely, if we take the interaction effect into consideration, it seems that new businesses will have a much better chance of survival if they are located in areas with a low spatial concentration of competitors but a high spatial concentration of related industries. This evidence is in line with some results obtained in the economic literature on firm survival and agglomeration economies. For example, just to mention one of the most recent studies, Howell et al. (2018) find a large and positive effect of related industries on firm survival in China. Also, in the case of tourism, some studies find similar results but on aggregate data (Capone & Boix, 2008; Lazzeretti & Capone, 2009; Yang & Fik, 2014). For example, Capone and Boix (2008) find that the growth rate of local tourism depends strongly on the chain of economic activities directly related to tourism production (defined by them as 'tourist filière') rather than natural resource endowment.

Hence, we can conclude that new firms in the accommodation industry may benefit in terms of survival probability from the presence of related industries but only when the spatial concentration of competitors is low. In other words, the 'congestion effect' related to *Localization* nullifies the benefits of spatial concentration of related industries.

As regards *Urbanization*, we do not find any significant effect. However, it is worth mentioning that the municipality shared frailty does have a significant effect. This means that new firms located in the same municipality are affected by some unobserved factors that influence the risk of failure.

In line with the aim of our research, we extended the basic specification of model 2 in order to take into account the effect on survival of geography and firm characteristics. We considered the following variables: *Size, Distance from the coast, Altitude* and *Accommodation* (see model 3). The modelling selection has been done on the basis of the goodness of fit and whether the hypothesis of proportional hazards assumption is respected. The goodness of fit is here measured by means of the Cox–Snell pseudo where its maximum limit is not absolute and depends on the sample; in our case it is equal to 0.93 for both models, indicating a quite good fit (Cox & Snell, 1989). The Grambsch–Therneau (PH) test was also performed to assess the hypothesis of proportional hazards assumption, which is only respected in model 3. Hence, this model has to be preferred because the goodness of fit does not decrease, and the preconditioned assumptions of the Cox proportional hazard model are respected.

Estimates of model 3, first of all, confirm the results from model 2. We also observe that new medium and large size firms are better placed to reduce the risk of failure: a higher turnover in the first year is important to the probability of survival.

As far as geographical location is concerned, we observe that the risk of market exit increases significantly when the business is > 2 km from the coast, but this evidence must be carefully interpreted, since the estimate is weakly significant. Nevertheless, this result does demonstrate how important coastal tourism is to Sicily. That seems to be confirmed by the coefficient of altitude, which is not statistically significant. These results might indicate the presence of the 'funnelling effect' (Pearce, 1996) whereby tourists remain concentrated along the seafront and miss the opportunity of discovering the other tourist products that Sicily has to offer, such as mountain or cultural experiences. These results, however, should be interpreted with caution due to the nature of the variables used. In fact, distance from the coast is a more accurate measure than altitude, since the former is obtained from georeferenced microdata while the latter is based on the altitude of the municipalities where firms are located.

Lastly, we do not find any significant effect in the comparison between hotels and other types of accommodation, while the result on the municipality shared frailty is confirmed.

Robustness check

After estimation, some unobserved spatial dependence could still affect our residuals. We check this eventuality by means of the variogram described below (Diggle & Ribeiro, 2007).



Figure 4. Monte Carlo envelopes for the variogram of the model 2 martingale residuals (dashed lines) and empirical variogram of residuals (circles).

For the martingale residuals of the Cox proportional-hazards model, the empirical variogram ordinates are the quantities

$$v_{ij}=\frac{1}{2}(r_i-r_j)^2,$$

where r_i and r_j are the martingale residuals corresponding to the firms at the locations $s_{b,i}$ and $s_{b,j}$, respectively. A plot of v_{ij} against the corresponding distance d_{ij} compared with the envelope of empirical variograms computed from random permutations of the residuals, holding their locations fixed, enables us to formally detect spatial autocorrelation. Therefore, we obtained variogram envelopes from 999 independent random permutations of the martingale model residuals with values averaged within distance bands. These Monte Carlo simulation envelopes represent confidence bands for the null hypothesis of no spatial dependence among the model residuals. Consequently, if all the empirical variogram ordinates are within the bands, we conclude that there is no significant residual spatial dependence.

Figure 4 depicts the empirical variogram of our estimated model residuals along with the confidence bands. For ease of interpretation, the v_{ij} are averaged within distance ranges. The plot shows that the empirical variogram ordinates are within the simulation envelopes, thus indicating that there is no significant spatial dependence amongst the model residuals. Therefore, we can say that including variables of a spatial nature in the model enabled us to fully control the estimates for potential spatial dependence.

CONCLUSIONS

Recently, the tourism literature is extending its lines of research towards entrepreneurship and industrial studies in order to investigate the performance of firms operating in this sector. In fact, over the past few years, tourism has had a significant economic impact with the potential to reduce the gap between peripheral and central geographical areas. However, this new stream of literature is still quite limited with few studies dealing with the topic of firm survival in the tourism industry and only two studies also taking into account the role of geography (Falk, 2013; Gémar et al., 2016).

To fill these gaps in the literature, our study makes a contribution from several different perspectives. First, we suggest an original framework of analysis in tourism economics literature, where agglomeration economies and geography are associated with the ability of accommodation businesses to survive in the market. Second, we provide novel evidence on the accommodation industry in the insular region of Sicily in Italy. Our empirical study could represent a starting point for a potential stream of new studies on the role of geography in the tourism industry in the specific case of insular and/or peripheral regions. The importance of this aspect is evident if one thinks that tourism will dramatically change as a consequence of the recent COVID-19 pandemic. Indeed, it is reasonable to expect that insular and peripheral areas could over time become more attractive for tourists than central and more densely populated areas.

Exploiting microdata at the census level on 660 Sicilian firms in the accommodation industry born during the period 2010–14, we initially find a non-homogeneous spatial distribution of firms for both births and deaths. Coastal areas and the biggest cities typically have large numbers of births and deaths. Some positive spatial patterns, that is, more births than deaths, are also to be found along the coast in places some distance from large urban areas.

From the regression analysis, we find that new firms in the accommodation industry suffer considerably from a 'congestion effect', as measured by the spatial concentration of firms in the same industry. We find a sort of barrier to entry in the presence of a spatial concentration of firms specialized in the accommodation industry. However, we do find beneficial effects on survival when new firms are located near to spatial concentrations of firms operating in related industries.

Moreover, we find that turnover at birth is an important factor in longer term survival; we can say that 'to be too small at birth' is risky. Besides, risk of market exit appears to be dependent on the proximity of desirable locations, especially at the coast. We find empirical evidence that firms should not overstep the threshold of 2 km from the coast to reduce the risk of failure. Clearly, the decision of firm localization is likely co-determined both by tourist demand and the attractiveness of destinations.

Our study does have some limitations, although most of these depend on the availability of the data. For example, we do not have data on firm characteristics such as whether the hotel is part of a chain or variables to measure management quality. In other words, we would need additional information to be sure that by controlling for firm characteristics we are able to isolate the contribution of external factors such as that of agglomeration economies or, more generally, geographical features.

ACKNOWLEDGMENTS

The authors are especially grateful to Giuseppe Espa for discussing with them the research idea and encouraging them to develop it. They thank all those attending the 49th Scientific Meeting of the Italian Statistical Society, especially Giuseppe Arbia and Roberto Patuelli, for their suggestions on a previous version of the article. The authors also thank two anonymous referees and the Editor-in-chief, Paul Elhorst, for their useful and constructive comments.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

NOTES

¹ However, they consider an extensive definition of localization economies including inside also the related industries, that is, the diversification.

 2 We referred to the following NACE codes: 5510 (hotels); 5520 (holiday homes, cottages, youth hostels, etc.); 5530 (campgrounds, recreational camps, space for recreational vehicles); and 5590 (student residences, worker hostels, rooming and boarding houses, etc.).

³ We excluded from the analysis firms located in the smaller islands off Sicily (the Eolian, Egadi and Pelagian islands). Moreover, we observe firms that started in any calendar year within the period 2010–14 and whether or not they survived until the end of the period. Therefore, the data set consists of right-censored longitudinal data. However, according to the typical approach of survival analysis, the dependent variable in this study is the survival time in each year (i.e., 1, 2, 3, etc. years), regardless of the calendar year of the company's birth and death.

⁴ See Appendix A in the supplemental data online for more details.

⁵ Specifically, we consider the following industries: restaurants and mobile food service activities (5610); beverage service activities (5630); renting and leasing of motor vehicles (7711, 7721); travel agency and tour operator activities (7911, 7912); other reservation service and related activities (7990); creative, arts and entertainment activities (9004); libraries, archives, museums and other cultural activities (9101–9104); sport activities (9319); and amusement and recreation activities (9321, 9329).

⁶ Since we exploit data at establishment level, there is no ambiguity about the attribution of annual turnover. However, using data at the registered office level could generate some bias in exploring the distribution of economic activities across space. Unfortunately, we do not know if firms belong to a group (e.g., a hotel chain), thus we cannot capture those potential economies of scale.

⁷ In place of distance from the coast, we could use distance from the beach since some parts of the coast could simply be rocks and therefore less attractive to tourists. However, it would be difficult to find an objective measure of distance from the beach, and this could significantly affect our results. The Sicilian coast is quite accessible almost everywhere and accommodation businesses are mostly located in all these accessible areas, but at different distances from the coast. In other words, different distances correspond in our analysis to different investments by firms in terms of location choice.

⁸ Other categories (e.g., from 2 to 3 km and so on) were considered, but no significant results emerged. Thus, we decided to stop at 2 km since it seems to be the significant threshold of the risk of failure.

⁹ All the firms observed remain in the market for at least one year after their entry.

¹⁰ See Tables C1–C7 in Appendix C in the supplemental data online for details on survival probabilities estimates.

¹¹ All statistical analyses were performed using the 'survival' package (Therneau, 2015) in the R statistical environment (R Core Team, 2018).

¹² From a macroeconomic perspective on the US economy, Saito and Wu (2015) find that employment growth may be negatively affected by employment density, suggesting an offset effect of congestion on localization.

ORCID

Davide Piacentino D http://orcid.org/0000-0001-6252-1264 Martina Aronica D http://orcid.org/0000-0002-8597-5215 Diego Giuliani D http://orcid.org/0000-0002-7198-6714 Andrea Mazzitelli D http://orcid.org/0000-0001-5690-0815 Maria Francesca Cracolici D http://orcid.org/0000-0002-9417-0937

REFERENCES

- Alsos, G. A., Eide, D., & Madsen, E. L. (2014). Handbook of research on innovation in tourism industries. In G. A. Alsos, D. Eide, & E. L. Madsen (Eds.), *Introduction: Innovation in tourism industries* (pp. 1–26). Edward Elgar.
- Arbia, G., Cella, P., Espa, G., & Giuliani, D. (2015). A micro spatial analysis of firm demography: The case of food stores in the area of Trento (Italy). *Empirical Economics*, 48(3), 923–937. https://doi.org/10.1007/ s00181-014-0834-6
- Audretsch, D. B., & Fritsch, M. (1994). The geography of firm births in Germany. *Regional Studies*, 28(4), 359– 365. https://doi.org/10.1080/00343409412331348326
- Balaguer, J., & Cantavella-Jorda, M. (2002). Tourism as a long-run economic growth factor: The Spanish case. Applied Economics, 34(7), 877–884. https://doi.org/10.1080/00036840110058923
- Boschma, R., & Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography*, 85(3), 289–311. https://doi.org/10.1111/j.1944-8287.2009.01034.x
- Boschma, R., Minondo, A., & Navarro, M. (2012). Related variety and regional growth in Spain. Papers in Regional Science, 91(2), 241–256. https://doi.org/10.1111/j.1435-5957.2011.00387.x
- Boschma, R. A., & Wenting, R. (2007). The spatial evolution of the British automobile industry: Does location matter? *Industrial and Corporate Change*, 16(2), 213–238. https://doi.org/10.1093/icc/dtm004
- Brau, R., Lanza, A., & Pigliaru, F. (2007). How fast are small tourism countries growing? Evidence from the data for 1980–2003. *Tourism Economics*, 13(4), 603–613. https://doi.org/10.5367/00000007782696104
- Brida, J. G., Cortes-Jimenez, I., & Pulina, M. (2014). Has the tourism-led growth hypothesis been validated? A literature review. *Current Issues in Tourism*, 19(5), 394–430. https://doi.org/10.1080/13683500.2013.868414
- Brouder, P., & Eriksson, R. H. (2013). Staying-power: What influences micro-firm survival in tourism. *Tourism Geographies*, 15(1), 125–144. https://doi.org/10.1080/14616688.2011.647326
- Capone, F., & Boix, R. (2008). Sources of growth and competitiveness of local tourist production systems: An application to Italy (1991–2001). *The Annals of Regional Science*, 42(1), 209–224. https://doi.org/10.1007/s00168-007-0133-7
- Chhetri, P., Corcoran, J., & Hall, C. M. (2008). Modelling the patterns and drivers of tourism related employment for south-east Queensland, Australia – A spatial econometric approach. *Tourism Recreation Research*, 33 (1), 25–38. https://doi.org/10.1080/02508281.2008.11081287
- Cox, D. R. (1972). Regression models and life-tables (with discussion). Journal of the Royal Statistical Society, Series B, 34, 187–220.
- Cox, D. R. (1975). Partial likelihood. Biometrika, 62(1975), 269-276. https://doi.org/10.1093/biomet/62.2.269
- Cox, D. R., & Snell, E. J. (1989). Analysis of binary data (2nd ed.). Chapman & Hall.
- Cracolici, M. F., & Nijkamp, P. (2008). The attractiveness and competitiveness of tourist destinations: A study of Southern Italian regions. *Tourism Management*, 30(3), 336–344. https://doi.org/10.1016/j.tourman.2008.07.006
- Delgado, M., Porter, M. E., & Stern, S. (2010). Clusters and entrepreneurship. Journal of Economic Geography, 10 (4), 495–518. https://doi.org/10.1093/jeg/lbq010
- Diggle, P. J., & RibeiroJr., P. J. (2007). Model-based geostatistics. Springer.
- Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. Journal of the American Statistical Association, 72(359), 557–565. https://doi.org/10.1080/01621459.1977.10480613
- Estevão, C., & Ferreira, J. (2009). The tourism clusters role in regional development: Presenting a competitiveness conceptual model. Tourism Destination Development and Branding Conference Proceedings, Eilat, 127–139.
- Falk, M. (2013). A survival analysis of ski companies. *Tourism Management*, 36, 377–390. https://doi.org/10. 1016/j.tourman.2012.10.005
- Fotopoulos, G., & Louri, H. (2000). Location and survival of new entry. *Small Business Economics*, 14(4), 311-321. https://doi.org/10.1023/A:1008180522759
- Frenken, K., Cefis, E., & Stam, E. (2015). Industrial dynamics and clusters: A survey. *Regional Studies*, 49(1), 10–27. https://doi.org/10.1080/00343404.2014.904505

- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5), 685–697. http://doi.org/10.1080/00343400601120296
- Fritsch, M., & Schroeter, A. (2011). Why does the effect of new business formation differ across regions? Small Business Economics, 36(4), 383–400. https://doi.org/10.1007/s11187-009-9256-9
- Gémar, G., Moniche, L., & Morales, A. J. (2016). Survival analysis of the Spanish hotel industry. *Tourism Management*, 54, 428–438. https://doi.org/10.1016/j.tourman.2015.12.012
- Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in cities. Journal of Political Economy, 100(6), 1126–1152. https://doi.org/10.1086/261856
- Grambsch, P., & Therneau, T. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81(3), 515–526. https://doi.org/10.1093/biomet/81.3.515
- Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics*, 53(1), 1–28. https://doi.org/10. 1016/S0094-1190(02)00505-3
- Howell, A., He, C., Yang, R., & Fan, C. C. (2018). Agglomeration (un)-related variety and new firm survival in China: Do local subsidies matter? *Papers in Regional Science*, 97(3), 485–500. https://doi.org/10.1111/pirs. 12269
- Jacobs, J. (1969). The economy of cities. Vintage.
- Kaplan, E. L., & Maier, P. (1985). Nonparametric estimation from incomplete observations. Journal of the American Statistical Association, 53(282), 457–481. https://doi.org/10.1080/01621459.1958.10501452
- Lazzeretti, L., & Capone, F. (2009). Spatial spillovers and employment dynamics in local tourist systems in Italy (1991–2001). European Planning Studies, 17(11), 1665–1683. https://doi.org/10.1080/09654310903230616
- Li, K. X., Jin, M., & Shi, W. (2018). Tourism as an important impetus to promoting economic growth: A critical review. *Tourism Management Perspectives*, 26, 135–142. https://doi.org/10.1016/j.tmp.2017.10.002
- Liu, J., Nijkamp, P., & Lin, D. (2017). Urban-rural Imbalance and tourism-led growth in China. Annals of Tourism Research, 64, 24–36. https://doi.org/10.1016/j.annals.2017.02.005
- Nakamura, R. (1985). Agglomeration economies in urban manufacturing industries: A case of Japanese cities. Journal of Urban Economics, 17(1), 108–124. https://doi.org/10.1016/0094-1190(85)90040-3
- Neves Sequeira, T., & Maçãs Nunes, P. (2011). Does tourism influence economic growth? A dynamic panel data approach. *Applied Economics*, 40(18), 2431–2441. https://doi.org/10.1080/00036840600949520
- Nooteboom, B. (2000). Learning and innovation in organizations and economies. OUP.
- Pearce, D. G. (1996). Tourism today: A geographical analysis. Longman Scientific and Technical.
- Peiró-Signes, A., Segarra-Oña, M. D. V., Miret-Pastor, L., & Verma, R. (2015). The effect of tourism clusters on US hotel performance. *Cornell Hospitality Quarterly*, 56(2), 155–167. https://doi.org/10.1177/ 1938965514557354
- Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., Mantel, N., McPherson, K., Peto, J., & Smith, P. G. (1977). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. II. Analysis and examples. *British Journal of Cancer*, 35(1), 1–39. https://doi.org/10. 1038/bjc.1977.1
- Piacentino, D., Espa, D., Filipponi, D., & Giuliani, D. (2017). Firm demography and regional development: Evidence from Italy. *Growth & Change*, 48(3), 359–389. https://doi.org/10.1111/grow.12172
- Porter, M. (2003). The economic performance of regions. *Regional Studies*, 37(6–7), 549–578. https://doi.org/10. 1080/0034340032000108688
- R Core Team. (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
- Renski, H. (2011). External economies of localization, urbanization and industrial diversity and new firm survival. *Papers in Regional Science*, *90*(3), 473–502. https://doi.org/10.1111/j.1435-5957.2010.00325.x
- Rosenthal, S. S., & Strange, W. C. (2001). The determinants of agglomeration. *Journal of Urban Economics*, 50(2), 191–229. https://doi.org/10.1006/juec.2001.2230
- Rosenthal, S. S., & Strange, W. C. (2004). Evidence on the nature and sources of agglomeration economies. In V. Henderson & J. F. Thisse (Eds.), *Handbook of regional and urban economics* (Vol. 4, pp. 2119–2171). Elsevier.

- Saito, H., & Wu, J. (2015). Agglomeration, congestion, and U.S. Regional disparities in employment growth. Journal of Regional Science, 56(1), 53-71. https://doi.org/10.1111/jors.12204
- Santarelli, E. (1998). Start-up size and post-entry performance: The case of tourism services in Italy. Applied Economics, 30(2), 157–163. https://doi.org/10.1080/000368498325967
- Segarra-Oña, M. D. V., Miret-Pastor, L. G., Peiro-Signes, A., & Verma, R. (2012). The effects of localization on economic performance: Analysis of Spanish tourism clusters. *European Planning Studies*, 20(8), 1319–1334. https://doi.org/10.1080/09654313.2012.680586
- Solvoll, S., Alsos, G. A., & Bulanova, O. (2015). Tourism entrepreneurship review and future directions. Scandinavian Journal of Hospitality and Tourism, 15(1), 120–137. https://doi.org/10.1080/15022250.2015. 1065592
- Therneau, T. (2015). A package for survival analysis in S. version 2.38, https://CRAN.R-project.org/package= survival.
- Yang, Y., & Fik, T. (2014). Spatial effects in regional tourism growth. Annals of Tourism Research, 46, 144–162. https://doi.org/10.1016/j.annals.2014.03.007