University of Reading

Location and survival of MNEs' subsidiaries: Agglomeration and heterogeneity of firms

Article

Accepted Version

Mariotti, S., Mosconi, R. and Piscitello, L. (2019) Location and survival of MNEs' subsidiaries: Agglomeration and heterogeneity of firms. Strategic Management Journal. ISSN 1097-0266 doi: 10.1002/smj.3081 Available at https://centaur.reading.ac.uk/86189/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1002/smj.3081

Publisher: Wiley

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

University of Reading

Central Archive at the University of Reading

Reading's research outputs online

Accepted Article

LOCATION AND SURVIVAL OF MNEs' SUBSIDIARIES: AGGLOMERATION AND HETEROGENEITY OF FIRMS

Sergio Mariotti
DIG-Politecnico di Milano
Email: sergio.mariotti@polimi.it
Via Lambruschini 4/b – 20156 Milano

Rocco Mosconi
DIG-Politecnico di Milano
Email: rocco.mosconi@polimi.it
Via Lambruschini 4/b – 20156 Milano

Lucia Piscitello (corresponding author)
Henley Business School & DIG-Politecnico di Milano
ORCID iD: 0000-0002-4817-315X
Email: lucia.piscitello@henley.ac.uk; lucia.piscitello@polimi.it
Henley Business School – University of Reading: Whiteknights, Reading RG6 6UD, UK
DIG-Politecnico di Milano: Via Lambruschini 4/b – 20156 Milano

Running head: Location and survival of MNEs' subsidiaries
Keywords: Agglomeration, Location, Heterogeneity, Survival, Foreign subsidiaries

LOCATION AND SURVIVAL OF FOREIGN MNEs' SUBSIDIARIES: AGGLOMERATION AND HETEROGENEITY OF FIRMS

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/smj.3081

This article is protected by copyright. All rights reserved.

Abstract

Research summary

Our study provides a quasi–replication of Shaver and Flyer (2000), which was among the first studies that challenged the positive role of agglomeration in determining companies' location choice and performances, thus changing the way management scholars view companies' attitude towards agglomeration forces.

We employ the same research design, specification and tests, and a different population, to discuss the generalizability of the original study. Building on the framework of Shaver and Flyer (2000), our findings offer intriguing new empirical evidence highlighting the importance of the differential between entering foreign firms and host country firms as a crucial condition in understanding agglomeration forces and adverse selection mechanisms.

Managerial summary

Our exercise confirms that agglomeration forces act differently on stronger *versus* weaker multinational enterprises (MNEs). However, we find that stronger MNEs tend to avoid location in highly specialized areas when they are afraid of knowledge leakages towards host country–based rivals that have enough absorptive capacity to benefit and improve their competitive advantages.

Managerial implications are quite relevant. Indeed, when MNEs avoid co–location in highly specialized areas, they also limit their own access to local knowledge and other agglomeration economies, such as supply networks and qualified workforce. Thus, MNEs managers need to design and implement devices that, on the one hand prevent local leakages of their knowledge and, on the other, do not hinder their access to local unique knowledge and resources.

INTRODUCTION

The role of agglomeration forces, i.e., the relationship between the characteristics of the local context and the behavior of firms is a critical issue, not only in strategic management. From Marshall (1890) onwards, economists and geographers have assessed — both theoretically and empirically — the importance of agglomeration in explaining the geographical distribution of activities and firms (see the reviews by Combes & Gobillon, 2015; and Rosenthal & Strange, 2004). Agglomeration economies concern the formation of a specialized local labor market, the vertical relationships along the value chain, and the emergence of industry specific spillovers. Thus, agglomeration effects refer to the mechanisms that make it more productive to cluster with other actors in the same area: reduced transportation and associated operational costs, and reduced costs of moving people across space, thus allowing labor market pooling and efficient allocation of labor

Accepted Article

This article is protected by copyright. All rights reserved.

(Puga, 2010), as well as face–to–face contacts, flows of ideas, and knowledge spillovers (for a collection on agglomeration economics, see Glaeser, 2010).

Along this line, business analysts and management scholars have been studying companies' location strategies correlating them with concentration of activities. In other words, firms locate their activities in agglomerated clusters, as long as they can exploit agglomeration externalities and reduce their costs (Hennart & Park, 1994; Mariotti & Piscitello, 1995).

This was the dominant *credo* about firms' location behavior until the publication of Shaver and Flyer (2000), henceforth SF, which provided a critical discussion challenging the attraction role of local agglomeration economies. On the one hand co–locating with other actors provide firms with agglomeration benefits, on the other hand it might also generate costs, previously understated by the literature, and the net balance is not always obvious. SF argued that externalities stemming from geographic clustering are not automatically positive and that in certain circumstances firms gain more from avoiding agglomeration. Specifically, "firms vary in the net benefit they receive from agglomeration economies. For example, firms with the best technologies, human capital, training programs, suppliers or distributors will minimally benefit from access to competitors' technologies, human capital, training programs, suppliers or distributors" (SF, 2000: 1175), while the opposite will be true for weaker firms.

In other words, firms' heterogeneity matters: 'good technology firms' may prefer to remain isolated, as they fear to dissipate their competitive advantages through the generation of externalities that benefit local competitors, thus seeing agglomeration economies as a centrifugal her than a centripetal force. Conversely, 'weak technology firms' would be more inclined to agglomerate as they bet on a favourable balance between positive and negative externalities they can access through agglomeration, thus inducing an adverse selection effect, according to which only weak firms will be keen on agglomerating.

Reflecting its overall impact, the SF study is well cited, with more than 960 Google Scholar citations as in June 2019, with an average number of citations equal to 67.6 per year in the last decade (2009–2018), and 380 Web of Science citations (with an increasing trend in the recent years). Since then, while previous literature used to consider agglomeration effects generally

Accepted Article

This article is protected by copyright. All rights reserved.

beneficial, at least until offset by congestion effects and pecuniary diseconomies (e.g., Richardson, 1995; Rosenthal & Strange, 2004), several studies in different fields (international business, global strategy, regional science and economic geography) have increasingly emphasized the ambivalent role of agglomeration in relation to the firm–level attributes and differential abilities to access and generate knowledge (e.g., Alcácer & Chung, 2007; Alcácer, Dezso, & Zhao, 2013; Li & Bathelt, 2018; Mariotti, Piscitello, & Elia, 2010; McCann & Folta, 2011; Rigby & Brown, 2015).

However, and regardless of the country, spatial concentration of activities is an undisputable fact of the economic landscape (McCann & Folta, 2008). There is abundant evidence that the urge to agglomerate remains strong (Glaeser, 2010), and that even companies with higher knowledge stocks (i.e., those companies that, according to SF should avoid agglomeration) show a tendency to, and benefit from, agglomeration (e.g., Goerzen, Asmussen, & Nielsen, 2013; McCann & Folta, 2011). Thus, the conclusions from the debate are far from clear due to studies either confirming (along the path started by SF) or denying adverse selection in location behavior (although confirming asymmetric but positive effects, as in, e.g., Hervas–Oliver, Sempere–Ripoll, Alvarado, & Estelles–Miguel, 2018).

We suggest that a re–examination of the relationships and results of the SF study may offer important insights to strategy research with respect to the role of agglomeration economies in relation to the differences in capabilities and competitiveness of entering foreign firms *versus* host country firms.

As for the design of our study, we adopt a quasi–replication (see Bettis, Helfat, & Shaver, 2016), i.e., we employ the same specification and tests adopted in the original study in a different context. The sample context of their study was the US, and adopted the state as the boundary for agglomeration benefits, mainly because "previous studies of foreign investment location show that the state is an important spatial unit that foreign investors consider when making direct investment location choices" (SF, 2000: 1181). Namely, they consider 48 mainland states. As we aim to assess whether the geographical and temporal context may be an important condition driving their results, we consider foreign firms' location choices into the twenty administrative regions in Italy as the unit of location analysis. Although Italian regions are smaller than the US states (the average

Accepted Article

This article is protected by copyright. All rights reserved.

region's population is about three million, *versus* six million for US states), they present autonomous administrative boundaries somehow similar to the US states, and a certain degree of decisional autonomy (e.g., in the attraction policies/measures for FDI, that is mostly relevant for the study of MNEs' location choices).

This allows us to quasi–replicate their study on a cross section of 90 greenfield entries (they had 101 events), observed in a single period, and study their survival after eight years, thus concluding whether the original findings are consistent and generalizable in other contexts. Firstly, by conducting a replication of the original work, we provide additional empirical evidence that confirms the positive relationships between MNEs' locational preferences and agglomeration economies associated to the high proportion of companies in the same industry. Second, by repeating their analysis about the survival of foreign companies, we test the reproducibility of their result on adverse selection, building additional information that may corroborate or weaken their results.

Our quasi–replication offers a number of important contributions to the managerial and economic literature on firms' location choice, behavior and performances. Furthermore, we build on SF (2000) study by considering how these relationships may change in a different empirical context. In doing so, our study offers insights into the generalisability of the original research, and adds to the knowledge on MNEs' locational behavior.

Specifically, by finding evidence on adverse selection mechanisms once allowing for the character of the local context, we extend the work of SF (2000). In fact, we claim that adverse selection happens only when strong foreign MNEs face local companies that are good enough to absorb technological spillovers and benefit from them while, when local companies are weaker, foreign investors are not afraid of agglomeration. Although this result seems at odds with SF's result (according to which strong foreign MNEs always prefer avoiding agglomeration), we believe the results allow to reveal interesting contingencies and boundary conditions behind adverse selection. In fact, MNEs entering the US in the late '80s were largely Japanese MNEs that deliberately avoided agglomeration, while preferring location in *greenfield sites*, that "provide access to fresh recruits — rural labor with little or no experience in manufacturing industry, who can be socialised to Japanese production requirements" and "enable firms to exert significant

Accepted Article

This article is protected by copyright. All rights reserved.

control over their workforce, in keeping with the traditional Japanese practice of erecting barriers to labor mobility" (Kenney & Florida, 1992: 24). In other words, location of Japanese *manufacturing transplants* was mainly driven by forces that differ from the conventional agglomeration economies: "traditional approaches to location theory inadequately explain the transplant complex" (Kenney & Florida, 1992: 32). We thus feel that given the weight of Japanese investments in the SF sample (about the 70percent), their results are likely to reflect the mentioned Japanese behavior in the US in the late '80s.

Our results about the role of the heterogeneity in capabilities/competitiveness of companies in their agglomerative behavior are in line with other studies inspired by SF (2000), i.e., Alcacer and Chung (2007). Namely, when taking into account local excellences, i.e., the presence of indigenous firms that are capable to effectively absorb knowledge leakages and increase their competitive advantage, we find that MNEs differ in their location choice, as they differently evaluate the balance between the positive agglomeration externalities and the negative externalities stemming from the risk of knowledge leakages towards local firms. On the one hand, stronger foreign firms agglomerate less in highly specialized areas populated by excellent local competitors. In fact, in comparison with the latter, stronger MNEs may suffer from disadvantages due to their *liability of outsidership* (Johanson & Vahle, 2009). Endogenous absorptive capacity being equal, local competitors benefit from being well rooted in the local network of relationships that help them in monitoring, detecting and accessing external knowledge generated by co–located firms, while MNEs are outsiders, not (or scarcely) embedded in the network. Thus, "outsidership, in relation to the relevant network, is the root of uncertainty" (Johnson & Vahlne, 2009: 1411) that makes foreign MNEs value costs (associated with loosing knowledge towards local companies) as definitely higher that the more uncertain benefits (associated with learning from the local context). On the other hand, weaker MNEs are more likely to perceive a positive balance between the agglomeration externalities and the outflows of knowledge and technologies, i.e., co–location always offer something good to gain from the firms in the host location and nothing to lose. In other words, while foreign MNEs tend in general to agglomerate, the adverse selection mechanism proposed by SF does emerge for stronger MNEs facing stronger local competitors, i.e., when taking into account the interaction between the type of MNEs and the heterogeneity of the local context.

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

DATA SOURCES AND VARIABLES DEFINITION

The research design suggested by SF adopts three stages: (i) in the first, they model foreign firms' location choice as a function of agglomeration externalities stemming from "the effect of similar industry activity" (SF, 2000: 1183). Their hypothesis, in line with most of the previous empirical literature, is that many firms will cluster, due to agglomeration and specialization economies; (ii) in the second stage, they model the survival of local establishments as a proxy for the firms' performance; namely, due to the hypothesized adverse selection, i.e., to the agglomeration bias leading only weaker firms to agglomerate, they expect (and find) a negative relationship between agglomeration economies and plant survival; (iii) they finally re–run the location choice models of stage (i) by taking into account the firm's heterogeneity in order to show the different approach of companies towards agglomeration.

Our empirical exercises replicate these steps. As in SF, we use the same data set in both location choice and survival models.

Data sources and sample

Our data come from the database Reprint–Politecnico di Milano, which contains yearly information about the Italian affiliates of foreign firms and the location of their plants, since 1986[1] (for further details, see Cominotti, Mariotti, & Mutinelli, 1999; and Mariotti, Mutinelli, & Sansoucy, 2014). Our sample is constituted by foreign greenfield entries occurred in Italy in the period 2000–2002. Since SF considered one single year, ideally we would have considered investments in 2001, i.e., the year of the Italian Industry and Services Census. As the number of greenfield FDI per year is quite low, we considered a three–year period around 2001 (i.e., 2000–2002). This amounts to 113 greenfield investments; however, as in SF, to ensure variability of the dependent variable within the same country in the model for survival, we excluded from the analysis all entries from countries with one single investment, as well as all entries from countries where all investments survived or none of the investments survived. This reduced the sample to 90 greenfield investments (33 in 2000, 31 in 2001 and 26 in 2002), undertaken by 83 foreign MNEs (not dissimilar from SF, i.e., 101

[1] The database Reprint-Politecnico di Milano provide a census of all the Italian affiliates of foreign firms. However, in the spirit of the replication, we did consider only greenfield ones.

This article is protected by copyright. All rights reserved.

greenfield investments undertaken by 95 foreign MNEs). As far as the industrial breakdown, we consider 103 three–digit Nace Rev. 2 sectors[2] (Eurostat, 2008), 37 of which witnessed at least one of the 90 initiatives in our sample.

Table 1A in the Appendix illustrates some characteristics of the 20 Italian regions in terms of land area, population, and number of establishment.[3] It is worth noting the strong and positive correlation (0.91) between the number of manufacturing establishments and the density of establishments. This means that Italian regions accounting for a large share on manufacturing establishments tend to have a high density of establishments that mirrors a high degree of urbanization, rather than a different geographical size.[4] 15 regions received at least one of the 90 greenfield initiatives in our sample; 7 regions hosted three or less foreign investments, five regions four to ten, and three regions more than 10.[5]

Variables definition

In the spirit of a replication study, the variables used in this study are constructed following SF as closely as possible. Table 1 reports the detailed description of our variables *versus* the variables adopted in the SF paper.

[Insert table 1 approximately here]

Agglomeration measure

The starting point for our analysis of the location choice is the same measure of the territorial specialization stemming from agglomeration forces used in SF, i.e., the proportion of industry establishments that are in the region in which a foreign plant locates. Thus, the variable $PROPORTION_{sj}$ is built as follows:

[2] Nace is the Statistical classification of economic activities in the European Community, available at http://ec.europa.eu/eurostat/documents/3859598/5902521/KS-RA-07-015-EN.PDF.

[3] To account for the different levels of geography between our study and SF, notice that the average land area of the Italian regions is about 13 times smaller than the average land area of the US States. The ratio drops to two when considering population, so evidencing a significantly higher spatial density of activities in Italy.

[4] As an example, Lombardy and Veneto account for a large share of manufacturing establishments (21.1% and 11.3% respectively versus an average share of 1/20=5%), and their density is much higher than the average (5.2 and 3.6 establishments per squared km, while the average density is 1.9).

[5] Specifically, Lombardy hosted 29 investments, Emilia-Romagna 14, Piedmont 12.

This article is protected by copyright. All rights reserved.

$$PROPORTION_{sj} = \frac{N_{sj}}{\sum_j N_{sj}}$$

where N_{sj} is the number of establishments[6] in sector *s* in region *j*.

Survival

In the original study, the performance is measured by *SURVIVAL*, i.e., a binary variable taking the value of one if the new plant continues to be controlled by the investing firm at the end of the observation period (i.e., 2009), or 0 otherwise. Therefore, a plant that still exists at the end of the period, but it is owned by some other investor, is regarded as non–survivor. We adopt the same measure; survival occurs in 58 percent of cases, i.e., 52 entries out of 90.

Control variables

We match the same set of control variables from the original study. Specifically, we expect that other industry–, firm–specific factors may affect firms' location choices as well as their survival.

Industry specific factors

As in the original study we control for *INDUSTRY GROWTH*, i.e., the proportionate change in industry shipments over the period 2001to 2007 (source: ISTAT). Indeed, firms entering fast–growing sectors could have survival advantages because they benefit from favourable market conditions and less fierce competition.

SF includes in the survival analysis the four–firm concentration ratio (*C4*), i.e., the market share of the largest four firms, to take into account that in concentrated industries incumbent firms are more likely to retaliate against new entrants, which might reduce profitability and hence the survival probability of new initiatives. Unfortunately, this indicator is not published in Italy. However, in the spirit of replication, we tried to approximate it as closely as possible by relying on Ellison and Glaeser (1997: Appendix B). Since the only information concerning the firms' dimension (at the needed sectoral disaggregation level) available in the 2001 census is the number

[6] As in SF, we use establishment counts to define $PROPORTION_{sj}$. However, this is highly correlated (0.90) with the alternative measure based on employment.

Accepted Article

This article is protected by copyright. All rights reserved.

of employees, while there is no information about sales, we started from the distribution of Italian firms by number of employees. We assumed a lognormal distribution, and we estimated the parameters using the available data. Using the estimated parameters, we have simulated the number of employees for the firms belonging to the largest class (250 employees or more),[7] under the condition that the sum of the employees is equal to the observed sum. We have repeated the simulation 100 times, and each time we have computed $C4_i$ $(i = 1, \dots, 100)$ as the sum of the 4 largest simulated values divided by the total number of employees in the industry. We then obtained C4 as an average of the $C4_i$'s. The descriptive statistics for our C4, provided in Table 2, differ somewhat from SF's (for us: Mean 13.13, St. Dev. 12.25, min 0.55, max 76.81; in SF: Mean 45.34, St. Dev. 21.59, min 5, max 90). However, we believe that this is not due to using employees rather than sales, or to the methodology used to estimate the size of the largest firms, but rather to the very different structure of Italian and US manufacturing. In fact, according to 2001 Italian census, the share of firms with more than 250 employees on total manufacturing employment was 22.9 percent, whereas, according to Enterprise Statistics, the corresponding figure in US manufacturing in 1987 was 73.1 percent.

Furthermore, to account for industry conditions that may favor foreign plants' survival, we include the variable *IT–FOREIGN* that measures the share of firms with international operations. In SF, this is proxied by the percentage of US firms reporting foreign income or foreign taxes in 1987 (from the *Compustat PCplus* database). We have a quite similar information for Italian industries in 2001. Our proxy is the proportion of firms with international operations among Italian firms with more than 10 employees in the industry in 2001, so to exclude craft businesses (source: ISTAT and ICE, 2002).

Finally, as in SF, we control also for the industry geographic concentration (*GEOGRAPHIC_CONCENTRATION*) through the traditional agglomeration index suggested by

[7] In the few cases where the largest class has fewer than 4 firms, following Ellison and Glaeser (1997) - Appendix B, the firms within the second-last class (100-249 employees) have been discretely uniformly spread on a range centered on the mean, with its boundary at the closer of the two end points of the size range. Finally, the employees of largest firms obtained this way have been added to the employees of the firms in the largest class to compute C4.

Accepted Article

This article is protected by copyright. All rights reserved.

Ellison and Glaeser (1997). This index has been calculated by relying on the 2001 Italian Industry and Services Census.

Firm–specific factors

As firms are likely to enjoy survival advantages if they have already previous country–specific experience, we introduce the two following variables: (i) *SUBSIDIARIES*, i.e., the number of subsidiaries (considering manufacturing, sales and distribution) controlled by the foreign parent, in Italy, in 2000. (ii) *RELATED*, i.e., a dummy taking value one if the parent firm controlled operations in the same two–digit sector in Italy in 2000. Information comes from the Reprint–Politecnico di Milano database.

In order to allow for the strength of the foreign entrant, we built a dummy variable (*LARGE_ESTABLISHMENT*) that allows for the relative size of the foreign initiative *versus* the average size of the establishments in the industry. The source is again the 2001 Italian Industry and Services Census. As in SF, as a further proxy of the strength of the investing firm we also included the dummy variable *FINANCIAL DATA*, being one if financial data were available for the foreign parent in 2001, zero otherwise. Actually, the dummy equals one in 95.5 percent of the cases in our sample (in SF the percentage was 55 percent), which casts some doubts on the appropriateness of this variable as a proxy for the strength of the parent in our context.

We also included seven dummies variables for the investors' home countries.[8] As in SF, these dummies account for factors that may affect survival, e.g., cultural and geographical distances, as well as differences in exchange rate in the period considered.

Finally, we added the variable *YEAR*, referring to the entry year (2000=-1, 2001=0, 2002=1). That allows controlling for the fact that recent initiatives might have a higher probability to survive in 2009, since they have been at risk for a shorter period. Unreported results show that adopting a more flexible (non–parametric) functional form by replacing the variable *YEAR* by two dummies (one for 2000 and one for 2002) leaves the results absolutely unchanged. Since the covariates

[8] The number of entries by nationality of the parent firm is as follows: 26 from Germany, 21 from the US, 20 from France, nine from Switzerland, eight from the UK, four from Japan, two from China.

This article is protected by copyright. All rights reserved.

Accepted Article

entered in the survival model vary slowly over time, we believe that considering their value in 2001 irrespective of the entry year is a very reasonable approximation.

Table 2 presents descriptive statistics for, and correlations among, all the variables.

[Insert table 2 approximately here]

MODELS AND RESULTS

Location choice

SF adopt the statistical method generally used to investigate foreign investment location choice, i.e., the conditional logit (McFadden, 1974). As they note, this is the appropriate model specification for data where we possess information on the location choices that an entrant made, and information about region attributes for the chosen location and other locations in the choice set. Namely, this technique estimates how differences in location attributes affect the probability that an entrant will choose to locate in a region. Specifically, if Π_{ij} is defined as the expected profit for investment *i* in region *j*, the expected profitability of investing in a region can be represented by the following equation:

$$\Pi_{ij} = \beta' X_{ij} + \varepsilon_{ij}$$

where X_{ij} is a vector of observed regional characteristics for investment *i* in region *j*, and ε_{ij} is a random disturbance that is attributable to errors associated with imperfect perception and optimization by the decision–maker and unobservable characteristics that affect the profitability of locating in a given region.

As shown in McFadden (1974), under suitable assumptions on ε_{ij}, the probability that a region *j* results in the highest profitability for an investment can be represented by the following logit expression:

$$\Pr(Y_i = j) = \frac{\exp(\beta' X_{ij})}{\sum_{k=1}^{20} \exp(\beta' X_{ik})} \qquad j = 1, \ldots, 20$$

This article is protected by copyright. All rights reserved.

where Y_i is the location of the *i*–th investment. This function can be estimated using maximum likelihood techniques and the estimates of β can be used to test if various regional characteristics significantly affect the probability of choosing a greenfield investment location.

Replication

Table 3 presents results of the direct replication of SF on our dataset. The first column reports the location model run on the entire choice set (i.e., 20 regions), while the second and the third column report the models run on a reduced number of cases and on a restricted choice sets respectively. The second column restricts the analysis to greenfield initiatives in geographically concentrated industries considering as a choice set all twenty regions, while in the third column we introduce regional fixed effects in order to control for possible bias in the estimates due to omitted variables (labor costs, local taxation, government incentives, or other omitted attributes of the regions): as illustrated in SF, this requires to restrict the choice set to the regions in which we have observed at least one greenfield initiative in the period considered.

Re–examining the functional form

As in the original study, in all the models the variable *PROPORTION* comes out positive and highly significant, thus confirming that firms tend to cluster with other firms in the same industry. However, in order to take into account the theoretical prediction about decreasing marginal returns to agglomeration (e.g., Combes & Gobillon, 2015), we have reconsidered the functional form, exploring the potential benefit of log transforming the variable *PROPORTION*. In fact, while the agglomeration benefits appear already with low levels of agglomeration, on the contrary, high level of agglomeration are likely to be associated with congestion and diseconomies (e.g., Broersma & van Dijk, 2008) that would decrease the likelihood of location. The estimates for the location model in which the variable *PROPORTION* has been replaced by *ln(PROPORTION)* are reported in Table 3b.

[Insert tables 3 and 3b approximately here]

The results confirm the positive impact of agglomeration economies. However, it is interesting to observe that the model in column 1 of Table 3b has a much higher likelihood than the model in column 1 of Table 3, despite the two models explain the location probabilities using the same

This article is protected by copyright. All rights reserved.

Accepted Article

explanatory variable (although with different functional forms). As expected, the logarithmic specification strongly outperforms the linear one (the difference in the log–likelihood is 10.66, while the number of parameters is the same). It is interesting to illustrate the differences between the two models. Figure 1 displays the estimated probabilities for each of the twenty alternatives in each of the 90 cases (1,800 estimated probabilities) based on model 1 in Table 3 and model 1 in Table 3b. The figure shows that the linear specification results in a highly nonlinear probability curve, where regions with *PROPORTION*>0.25 would be chosen by foreign investors with a much higher probability than what the logarithmic specification would predict; the opposite when with *PROPORTION*<0.25. In the linear model the vast majority of the 1,800 probabilities are not very different from each other, and are very close to 1/20=5 percent, while only in a few situations the probability would be much higher. Conversely, the logarithmic model in Table 3b, column 1, displays an almost linear probability curve, where the highest estimated probability is around 50 percent; clearly, also with the logarithmic specification we obtain lower probabilities when *PROPORTION* is low, but the dependence from *PROPORTION* is more clear, and the probabilities are frequently different from 1/20.[9]

[Insert Figure 1 approximately here]

Starting from the model in Table 3b, column 1, we narrow the analysis to the 43 initiatives undertaken in industries above the median level of geographic concentration (see column 2 in Table 3b). As in the linear model, the coefficients are almost identical to those based on all 90 entries.[10]

Also in this case, in order to control for possible bias in the estimates due to omitted variables (labor costs, local taxation, government incentives, or other omitted attributes of the regions), we introduce regional fixed effects, which requires to restrict the choice set to the 15 regions in which we have observed at least one greenfield initiative in the period considered. The estimates (excluding the coefficients on the dummies) are reported in column 3 of Table 3b. The estimated

[9] In the linear model we have only 195 cases out of 1,800 where the location probability is below two percent, and only 115 cases where the location probability is above 10 percent; the corresponding figures for the logarithmic specification are 814 and 277 respectively, showing that the latter specification discriminates different situations much more accurately.

[10] SF argue in favour of a stronger impact of agglomeration on location choice in this case, although their empirical evidence in this sense is extremely weak.

Accepted Article

This article is protected by copyright. All rights reserved.

coefficient for *ln(SPECIALIZATION)* decreased due to the introduction of the fixed effects. However, the Hausman specification test comparing the estimates in column 1 and column 3 of Table 3b is equal to 2.93 (p–value=0.09 in the χ_1^2), which seems to exclude any bias on the coefficient of *ln(SPECIALIZATION)* due to omitted regional specific variables.[11]

Summarizing, the logarithmic specification has several advantages over the linear one: the likelihood increases, the estimates become more robust and stable, and the estimated probabilities are much more credible than those obtained with the linear model.[12]

Foreign investment survival

After confirming that, overall, foreign firms are more likely to locate in regions with higher agglomeration economies, we investigate whether that also affects the likelihood of survival. In particular, following SF, we run a binary logistic regression, and estimation results are reported in Table 4 and 4b.

Replication

Table 4 reports the results of the direct replication of SF on our dataset, with minor modifications related to some specificities of our study. We have introduced the variable *YEAR* (-1,0,1, where 0 means that the initiative started in 2001) to take into account that recent initiatives are more likely to survive in 2009, since they have been at risk for a shorter period.

Results show that, contrary to SF results, agglomeration economies associated to the proxy adopted by SF show a positive sign, although weakly significant. The evidence seems the same, although less significant, also when the sample is restricted to geographically concentrated industries (column 2). Actually, as in SF, few of the variables introduced in the model seem to have significant explanatory power for *SURVIVAL*.

[11] Notice that the Hausman test comparing columns 1 and 3 in Table 3 is equal to 11.34 (p-value=0.00 in the χ_1^2), indicating that the estimates in the linear model are way less robust.

[12] As an example, consider the sector 'Preparation and spinning of textile fibres': according to the linear specification in Table 3 column 1, the probability to locate in Tuscany (PROPORTION=0.521) would be 80.3 percent, whereas the logarithmic specification gives a more credible 50.4 percent (actually, none of the two observed investments in that sector located in Tuscany).

Accepted Article

This article is protected by copyright. All rights reserved.

Re–examining SF model for survival

Overall, as these results do not confirm the negative relationship hypothesized by SF, we decided to further investigate it. In particular, the weakly positive impact of *PROPORTION* on *SURVIVAL* in the Italian case strengthened our conviction about the need of better defining the heterogeneity of foreign firms in relation to the characteristics of the host country's firms. Indeed, the competitiveness of local firms significantly differ among industries, as a result of a different endowment of technologies, knowledge, and other competitive advantages accumulated and shared in the local contexts. Thus, the more the host country's industries exhibit excellence, the higher the likelihood that local firms would be equal or even superior to the stronger foreign firms in terms of both generation and absorption of knowledge spillovers (Mariotti *et al.*, 2010).

Under conditions of industry excellence, foreign firms' location behavior may strongly differ between weaker and stronger companies. Weaker MNEs will locate in an area specialized in that (excellent) industry, as they benefit from the Marshallian triad (specialized labor, suppliers and spillovers). In particular, as they do not fear to be an effective source of spillovers for local firms, their perception of the net benefit of co–location is overall positive. On the contrary, stronger foreign firms may suffer from potential leakages towards their local peers, which have the absorptive capacity to effectively capture spillovers generated by the foreign entries and to strengthen their competitive position both locally and internationally. Indeed, foreign MNEs may also locally access some specific technologies and knowledge, but the value of this absorption is uncertain, being the access hindered or strongly limited by the 'outsidership of MNEs', which are weakly embedded in the local business network. Thus, stronger MNEs conservatively evaluate a negative balance from the co–location with local firms.

As a result, the more the industries in which foreign firms operate are characterized by excellence at the host country level, the more the adverse selection mechanism proposed by SF is at work. In industries where local firms are excellent, stronger foreign firms tend to avoid the specialized areas, where their locally–based competitors are massively located, while weaker MNEs may find them neutral or even promising.

Accepted Article

This article is protected by copyright. All rights reserved.

In order to test such an argument, we rerun the survival model by introducing a measure of excellence of the local contexts, associated to their sectors of specialization.

We rely on the concept of international competitiveness of the national industry, as measured by the Revealed Comparative Advantage (RCA) index (Balassa, 1965). The index is defined as follows:

$$RCA_s = \frac{X_{s,Italy} \Big/ \sum_i X_{i,Italy}}{\sum_j X_{s,j} \Big/ \sum_i \sum_j X_{i,j}}$$

where $X_{i,j}$ is the export in sector *i* from country *j*. The numerator and denominator represent the share of a given sector *s* in the national export and world export, respectively.[13] This index varies around unity: values greater than 1 suggest that a country's sector is comparatively advantaged in the international arena, whereas values less than 1 witness a position of comparative disadvantage. Thus, *RCA* measures the international performance of country's sectors as a result of the distinctive competence mix of local firms.[14] The relevant data for the calculation of RCA values come from the UN Comtrade database.[15]

Starting from *RCA*, we create the dummy variable *EXCELLENCE* indicating whether RCA_s in the year before the entry is larger than 1.5. To transform RCA into a dummy we considered several thresholds. The results provided in the rest of the paper are essentially unchanged for any threshold

[13] It is easily seen that RCA_s is also equal to $\frac{X_{s,Italy} \Big/ \sum_j X_{s,j}}{\sum_i X_{i,Italy} \Big/ \sum_i \sum_j X_{i,j}}$, i.e., the ratio of Italy's share of world export in sector s over Italy's share of world export in manufacturing as a whole.

[14] The highest RCA values are recorded by 'made in Italy' sectors (homebuilding materials and furniture; leather, textiles and clothing; and machine tools and other machinery), which are the sectors in which Italy enjoys international leadership (De Benedictis, 2005). In contrast, the lowest RCA values result for high-tech sectors (electronics and electronic instruments, computers, communication equipment, basic chemicals, and aerospace); in these sectors, Italy lags far behind most other advanced countries.

[15] We collected data on export from Italy to all countries for 2001 in HS96 classification at the highest disaggregate level from https://comtrade.un.org/data, then converted from HS96 to four-digit ISIC 3.1 using the concordance tables found in WITS (https://wits.worldbank.org/product_concordance.html), and finally to Nace Rev. 2 three-digit classification.

Accepted Article

This article is protected by copyright. All rights reserved.

in the range from 1.3 (60th percentile) to 1.5 (70th percentile). Outside this range, the dummy becomes less significant. We believe that this makes sense: in fact, below 1.3 one gets too close to the median (and therefore the notion of excellence is lost), whereas moving in the opposite direction the number of cases would be too small, and the result become unstable and too dependent on the individual cases.

Table 4b reports the results obtained running the model both on the full sample (columns 1 and 1b) and on the restricted sample (column 2). Interestingly, we observe that the impact of agglomeration economies on the companies' survival drops very close to zero if the greenfield is in a sector of excellence. Following the hints in Hoetker (2007), the different impact of *PROPORTION* on the probability of survival in the sectors of excellence with respect to the sectors where Italian firms are weaker can be effectively illustrated in Figure 2, which is based on the estimates in Table 4b, column 1b. The plot shows how the survival probability varies as a function of *PROPORTION*, and the interaction with *EXCELLENCE*, considering a typical situation for the other variables (i.e., the dummies are set equal to the mode and the other variables are set equal to the median; the country is set to US). Modifying the reference situation, the plot would obviously change, as the probability corresponding to *PROPORTION*=0 would be different, but the main characteristics would be unchanged: the curve would remain almost constant if *EXCELLENCE*=1, while it would grow substantially when *EXCELLENCE*=0. Our interpretation is that when local companies possess those competencies and skills that make them capable to absorb knowledge leakages and spillovers from foreign firms, the best competitors decide to locate elsewhere: therefore, investors in that region and sector are likely to be weaker, leading to a lower survival probability. That seems therefore to provide a contingency where the local specialization in a given industry and the relevant externalities do negatively correlate with firms' survival, thus witnessing adverse selection processes going on in the location strategies. In order to test the latter, we propose to revise the location choice analyses and check whether adverse selection in highly specialized local areas is confirmed.

[Insert tables 4 and 4b approximately here]

[Insert Figure 2 approximately here]

Accepted Article

This article is protected by copyright. All rights reserved.

Revisiting the location choice analyses

We revisit the location choice analyses previously presented by adopting the same approach of SF. We test whether adverse selection is at work, and whether it is associated with different agglomerative behavior of stronger *versus* weaker MNEs in different industrial contexts. As in SF, we use the dummy variable *LARGE_ESTABLISHMENT* as a proxy of the strength of the entrant MNE.

Econometric estimates are reported in Tables 5 and 5b. Table 5 reports the results of the direct replication of SF on our dataset, which do not confirm their findings.[16] Stronger MNEs do not seem to agglomerate less than weaker ones. In fact, the interaction term *PROPORTION* × *LARGE_ESTABLISHMENT*, although negative, is not significant. The restrictions to geographically concentrated industries and to the choice set of regions in which at least one greenfield initiative is observed, do not show any different results (the two restrictions are reported in the second and third column, respectively).

[Insert tables 5 and 5b approximately here]

However, given the results from our survival model, we expect that stronger MNEs will be less inclined to agglomerate in specialized areas when local companies are more likely to be fierce competitors both at national and global level. Therefore, we re–estimated the models in which we added the terms *ln(PROPORTION)* × *EXCELLENCE* and *ln(PROPORTION)* × *LARGE_ESTABLISHMENT* × *EXCELLENCE*. The results are reported in Table 5b. According with SF (p. 1188), bigger establishments would generate externalities that disproportionately help their competitors; our econometric findings show that this is statistically significant only when the latter operate in host country's excellent industries (the estimated coefficient for the variable *ln(PROPORTION)* × *LARGE_ESTABLISHMENT* × *EXCELLENCE* is always negative, and significant). The plot in Figure 3 shows the estimated location probabilities as a function of *PROPORTION* as resulting from the model estimated in Table 5b, column 1b. It is interesting to observe that, because of the nonlinearity of the conditional logit model (see Hoetker, 2007), when

[16] Our starting point in this Section is the logarithmic specification given in Table 3b rather than the original SF linear specification. We also analysed the linear specification, and the results are identical: *LARGE_ESTABLISHMENT* has a negative but insignificant coefficient.

This article is protected by copyright. All rights reserved.

Accepted Article

considering the location of a large establishment in industries where Italian firms have high revealed competitive advantage, highly agglomerated regions are relatively less attractive, whereas less agglomerated regions are relatively more attractive.

[Insert Figure 3 approximately here]

Our interpretation is that in this case, local firms are more receptive, and able to absorb knowledge leakages and to gain competitive advantages. Thus, stronger MNEs will be less likely to agglomerate, or equivalently, weaker companies will be relatively more likely to agglomerate in those areas, as they bet on a positive balance in terms of benefits stemming from the interaction with local excellent competitors. As a consequence, and similarly to SF, MNEs' subsidiaries show a lower likelihood to succeed (i.e., survive) when locate in highly specialized areas in these industries.

We have also considered alternative ways to measure the strength of the entrant MNE. Following the same intuition leading to introduce the Revealed Comparative Advantage (RCA) index as a proxy for the international competitiveness of the Italian industry, we try to use the RCA index of the MNE's home country in the focal industry:

$$RCA_s^{HOME} = \frac{X_{s,HOME} \Big/ \sum_i X_{i,HOME}}{\sum_j X_{s,j} \Big/ \sum_i \sum_j X_{i,j}}$$

as an alternative proxy for the strength of the competitor, based on the idea that if the MNE's home country has a high RCA in the focal industry, we may argue that the entrant is likely to be strong. The UN Comtrade database used to compute *RCA* allow us to compute RCA_s^{HOME} as well for all 90 investments. We then transform RCA_s^{HOME} into a dummy using the same threshold, i.e. 1.5, adopted to define the variable *EXCELLENCE*. Hence, we create the variable *EXCEL_COMP*, which is equal to one if $RCA_s^{HOME} > 1.5$ in the relevant sector. Table 5c illustrates the estimates of the location model when *EXCEL_COMP* is used instead of *LARGE_ESTABLISHMENT*: results are very similar to those results illustrated in Table 5b in terms of sign and significance (the likelihood is slightly

This article is protected by copyright. All rights reserved.

better), thus supporting the idea that *EXCEL_COMP* may be regarded as related to the strength of the competitor.

[Insert table 5c approximately here]

Notice that in 47 of the 90 investments both *EXCEL_COMP* and *LARGE_ESTABLISHMENT* are equal to zero, in 13 cases only *EXCEL_COMP* is one, in 25 cases only *LARGE_ESTABLISHMENT* is one, while only in five cases both dummies are equal to one. The tetrachoric correlation is therefore very low (-0.11), and most of the large establishments are owned by MNE's whose home country is not specialized in the focal sector. It seems therefore that the two proxies capture different aspects of the strength of the competitor, and both seem to be associated to a tendency to avoid agglomeration. We therefore tried another proxy, i.e., the dummy variable LARGEorEXCEL, which takes value one when either *LARGE_ESTABLISHMENT* or *EXCEL_COMP* is equal to one (43 cases). The results obtained using this variable are illustrated in Table 5d, and they seem to marginally improve the previous ones (as the sign remains the same while the significance, as well as the likelihood, increase).

[Insert table 5d approximately here]

DISCUSSION

In order to reconsider SF results in the light of our findings, we assess how the two settings are different in terms of the foreign entrants' and local firms' competitiveness. We derived $RCAs$ (i.e., RCA of the United States in the focal sectors of each of the greenfield initiatives analyzed in SF) and RCA_s^{HOME} (i.e., RCA of the home country of the MNE in the focal sector s of each of the initiative) for the investments analysed in SF. Specifically, we recovered SF data from the Department of Commerce, International Trade Administration publication Foreign Direct Investment in the United States, 1987 Transactions[17]. As in SF, we have considered only greenfield entries, restricting the sample to include only entries in the manufacturing sector (SIC 2000–3999) and by parent firms from Canada, France, Germany, Japan, and the United Kingdom. This leads to 121 investments (out of 1328), whereas SF dataset includes only 101 entries. In the paper they

[17] See FRASER, https://fraser.stlouisfed.org/title/854/item/33325.

This article is protected by copyright. All rights reserved.

Accepted Article

specify that additional entries have been excluded because of "the inability to independently verify the entry, missing data regarding concentration ratios in the 1987 Census of Manufacturers, and joint venture entries by two or more foreign firms" (SF, 2000: 1180): we could not reproduce this step, and therefore our analysis is based on 121 investments, out of which 84 (69.4%) are from Japan, 12 (9.9%) from Germany, 10 (8.3%) from Canada, 9 (7.4%) from the UK and 6 (5.0%) from France (according to the figures in Table 1 of SF, in their sample 69 percent of the investments from Japan, 13 percent from Germany, four percent from Canada, nine percent from the UK and five percent from France). For each greenfield investment, the dataset includes the SIC four–digit code, which we have aggregated to three–digit (the 121 observed investments belong to 55 different three–digit SIC sectors).

We then moved to export data in order to construct RCA in 1987 for US and the competitors' home countries (Japan, Germany, Canada, UK and France) in the relevant SIC–3d sectors. We used the Comtrade database, downloading the data in HS classification as close to 1987 as possible. Export data became gradually available in HS classification: Japan and Germany were available from 1988, Canada from 1989, US from 1991, UK from 1993 and France from 1994. In 1994, the first year where all six countries are reported, the database covered only 96 countries, therefore we preferred to use 1996 data (120 countries) to have more stable and complete data.[18] Starting from the data on export from US to all countries for 1996 in HS96 classification at the highest disaggregate level, we converted from HS96 to three–digit SIC using the concordance tables found in WITS. The same has been done for the other five countries. Using these export data, we have derived the RCA indicators in the relevant sectors for US and for the home country of each investor. Actually, for three SIC sectors export data were not available; therefore, we had to drop all investments in SIC sectors 253 (1 case), 304 (1 case) and 307 (5 cases), which leaves 115 investments.

[18] It may be worth acknowledging that we also tried to use the Feenstra-Lipsey world trade flows data for 1987 (https://cid.econ.ucdavis.edu/wix.html, Feenstra, Lipsey, Deng, Ma, & Mo, 2005), and we converted them using the concordance between the SITC and the industry classification used in the outward investment surveys of the Bureau of Economic Analysis (BEA) of the U.S. Department of Commerce (Feenstra, 2000). However, as the BEA classification is too aggregate for our purpose, we decided to use Comtrade data. We are aware that constructing RCA indicators in 1996 instead of 1987 is suboptimal: however, given the purpose of the analysis, we believe that using 1996 data at the appropriate disaggregation level is to be preferred to using 1987 data on the BEA macro sectors.

Accepted Article

This article is protected by copyright. All rights reserved.

As previously done for our Italian 2001 data, we then converted RCA_s and RCA_s^{HOME} for SF investments into dummies: EXCELLENCE=1 indicates $RCA_s > 1.5$, while EXCEL_COMP=1 indicates $RCA_s^{HOME} > 1.5$. Tables 6 and 7 illustrate the distribution of EXCELLENCE and EXCEL_COMP for the Italian 2001 dataset and the US 1987 dataset.

[Insert tables 6 and 7 approximately here]

The diagonal elements in the two tables are very similar; the most remarkable difference is that, in the US 1987 dataset, there is a large share of cases where EXCELLENCE=0 and EXCEL_COMP=1 (32.5% *versus* 7.8% in Italy), whereas in few cases EXCELLENCE=1 and EXCEL_COMP=0 (9.6% *versus* 28.9% in Italy): in other words, in the US dataset there are many initiatives in sectors where the investor's home country is strong and the local competitors are weak, and just a few where the opposite occurs.

According to our findings (see Table 8, which is based on the results in Table 5c), we would therefore predict that strong competitors in US 1987 should agglomerate more than weak investors, since in most cases they should not fear local competitors, and they could benefit from agglomeration economies. Actually, this is at odds with SF finding: their Table 5 shows that strong investors (measured in their paper by LARGE ESTABLISHMENT=1) agglomerate less, a circumstance that we would predict only when EXCELLENCE=1.

[Insert table 8 approximately here]

Thus, to shed some light on this puzzling evidence, we checked the 37 cases (32.5%) in the US dataset where EXCELLENCE=0 and EXCEL_COMP=1. Table 9 illustrates the sectors and home country for such investments.

[Insert table 9 approximately here]

The table clearly shows that most of the investments are from Japan, and they are mainly in the motor vehicles and equipment sector and other related activities (metalworking and machinery). Thus, given that location behavior of Japanese *manufacturing transplants* in the US in the '80s, was mainly driven by forces that differ from the traditional agglomeration economies (Kenney & Florida, 1992), we believe that the conclusion on adverse selection affecting *always* stronger MNEs

Accepted Article

This article is protected by copyright. All rights reserved.

may have reflected that specific contingency and period in time. Instead, we find that adverse selection mechanisms are more likely to be at work when local companies are strong enough to absorb and redeploy MNEs' advantages.

CONCLUSIONS

Our replication study fully confirms that MNE subsidiaries locate in geographical areas with higher shares of industrial activities. Additionally, we also confirm the original intuition of SF about the heterogeneity in the agglomerative behavior of stronger versus weaker MNEs, as only the latter are willing to agglomerate, thus leading to adverse selection mechanisms. However, our results show that this evidence emerges only when some contingencies are taken into account. Specifically, we find that stronger MNEs are less keen to agglomerate when they bet on a negative balance between agglomeration benefits (associated with learning from the local context) and costs (associated with loosing knowledge towards local companies), namely in areas characterized by strong indigenous companies. In fact, while strong local competitors have high absorptive capacity, MNEs suffer from liability of outsidership (Johanson & Vahlne, 2009) that, ceteris paribus, make them less able to absorb locally.

Our evidence is in line with the main idea put forward by SF, confirming that "heterogeneity among firms can lead to asymmetric contributions and benefits from the agglomeration externality" (SF, 2000: 1191) and that, consequently, optimal location choices should account for both relative costs and advantages of agglomeration. However, results also suggests that MNEs' heterogeneity needs to be confronted with local firms' heterogeneity (e.g., Alcácer & Chung, 2007; 2014), and arguably other sources of heterogeneity relevant to the context (Cantwell, 2009).

Thus, the story we tell (concerning the motivations for the strongest MNCs to not agglomerate) is actually the same told by SF (2000). However, we contextualize the scenario by relying on the absorptive capacity of the local competitors, while their conclusion (i.e., that stronger MNEs are always less incline to agglomerate) seems to be crucially dependent on the peculiar behavior of — mainly Japanese — MNEs in the US at the end of the '80s. According to the relevant literature (Head, Ries, & Swenson, 1995; Kenney & Florida, 1992), Japanese manufacturing investment

Accepted Article

This article is protected by copyright. All rights reserved.

created a Japanese–style production complex (mainly in the automotive and related sectors) that traditional approaches to location theory were not capable to explain.

We believe this study has also relevant managerial implications. MNEs avoid co–location in more specialized areas when their proprietary technologies and knowledge can leak out, thus benefitting host country–based rivals. But in so doing, MNEs also limit their access to local knowledge and other agglomeration economies, such as supply networks and qualified workforce. In order to turn the negative cost–benefit balance into a positive one, MNEs managers need to carefully design and implement devices to prevent leakages of key knowledge to competitors. This is a very challenging task as devices traditionally used to limit knowledge leakages often hinder also knowledge absorption and transfer, being these processes inherently correlated (Easterby–Smith, Lyles, & Tsang, 2008; Jost & van der Velden, 2008).[19] However, investing managerial resources to improve contextually firm's appropriability and absorption may lead to high economic returns. With regard to location, the advantages stem from the exploitation of the full potential of agglomeration, as firms can establish their activities quite irrespectively of the local presence and competitive strength of rivals.

As for future research, we believe that the study of the relationship between agglomeration and location choices would benefit from a closer examination of heterogeneity of firms. Strengths and weaknesses of new entrants and indigenous companies might be captured along several dimensions (e.g., innovativeness, profitability, competitiveness, growth); MNEs' location choices may be influenced by experience and learning stemming both from own previous entries and from imitation of other foreign companies' location choices (e.g., Shaver, Mitchell, & Yeung, 1997; Belderbos, van Olffen, & Zou, 2011; Koçak & Özcan, 2013), thus impacting also their survival likelihood in each local context.

Additionally, *ex ante* selection mechanisms could be further investigated by taking into account that both agglomeration and survival of firms may be influenced by idiosyncratic characteristics of the locations (e.g., Hernandez, 2014), and/or vary according to the stage of the product lifecycle in

[19] As far as knowledge leakages, see the review by Durst *et al.* (2015) and some recent research on how firms can protect themselves from the risks of spillovers to rivals (Alnuaimi & George, 2016; Hernandez, Sanders, & Tuschke, 2015).

Accepted Article

This article is protected by copyright. All rights reserved.

an industry (Potter & Watts 2010; Neffke, Svensson, Boschma, Lundquist, & Olander, 2011; Wang, Madhok, & Xiao Li, 2014). However, while this may be an intriguing investigation, examining the temporal boundaries of these mechanisms would require long inter–temporal and cross–sectional series.[20]

Finally, as the inner mechanisms behind agglomeration depend on interactions among people, firms and social entities, and are therefore spatially localized (Combes & Gobillon, 2015), agglomeration effects present a rapid spatial decay (see Drucker, 2012; Duranton & Puga, 2004). Although the role of geographical proximity varies across industries and by types of agglomeration considered, spillovers attenuate rapidly across space (Rosenthal & Strange, 2001) and their "effects are rarely found to be significant beyond 100 kilometers" (Combes & Gobillon, 2015: 306).[21] Thus, analyses at a finer spatial granularity level may offer further evidence on the mechanisms behind the relationship between the firm's location choice and the perceived net benefits from local externalities and spillovers.

ACKNOWLEDGMENTS

The authors thank the Editor, Wilbur Chung, and three anonymous reviewers for their insightful and constructive comments throughout the process.

[20] For instance, Wang *et al.* (2014) observe about 100 years but in a single-industry setting.

[21] It is worth noting that in the 60 percent of studies on agglomeration effects surveyed by Drucker (2012), such a threshold is twenty kilometers or less, and in more than 80 percent of studies is 80 kilometers or less.

Accepted Article

This article is protected by copyright. All rights reserved.

REFERENCES

Alcácer, J., & Chung, W. (2007). Location strategies and knowledge spillovers. *Management Science, 53*, 760–776.

Alcácer, J., Dezso, C.L., & Zhao, M. (2013). Firm rivalry, knowledge accumulation, and MNE location choice. *Journal of International Business Studies, 44*, 504–520.

Alcácer, J., & Chung, W. (2014). Location strategies for agglomeration economies. *Strategic Management Journal, 35*, 1799–1761.

Alnuaimi, T., & George, G. (2016). Appropriability and the retrieval of knowledge after spillovers. *Strategic Management Journal, 37*, 1263–1279.

Balassa, B. (1965). Trade liberalization and revealed comparative advantage. *The Manchester School of Economic and Social Studies, 32*, 99–123.

Belderbos, R., van Olffen, W., & Zou, J. (2011). Generic and specific social learning mechanisms in foreign entry location choice. *Strategic Management Journal, 32*, 1309–1330.

Bettis, R.A., Helfat, C., & Shaver, M. (2016). The necessity, logic, and forms of replication. *Strategic Management Journal, 37*, 2193–2203.

Broersma, L., & van Dijk, J. (2008), The effect of congestion and agglomeration on multifactor productivity growth in Dutch regions. *Journal of Economic Geography, 8*, 181–209.

Cantwell, J.A. (2009). Location and the multinational enterprise. *Journal of International Business Studies, 40*, 35–41.

Combes, P.P., & Gobillon, L. (2015). The empirics of agglomeration economies. In G. Duranton, V. Henderson, & W. Strange (Eds.). *Handbook of Urban and Regional Economics* (pp. 247–348). Amsterdam: Elsevier.

Cominotti, R., Mariotti, S., & Mutinelli, M. (1999). *Multinational Italy.* Collana Documenti Cnel, 2nd Series, n.17. Rome: Cnel Press.

De Benedictis, L. (2005). Three decades of Italian comparative advantages. *The World Economy, 28*, 1679–1709.

Drucker, J.M. (2012). The spatial extent of agglomeration economies: evidence from three U.S. manufacturing industries. US Census Bureau Center for Economic Studies Paper No. CES–WP–12–01.

Duranton, G., & Puga, D. (2004). Micro–foundations of urban agglomeration economies. In J.V. Henderson & J–F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (pp. 2063–2117). Amsterdam: North–Holland.

Durst, S., Aggestam, L., & Aisenberg Ferenhof, H. (2015). Understanding knowledge leakage: A review of previous studies. *VINE Journal of Information and Knowledge Management Systems, 45*, 568–586.

Easterby–Smith, M., Lyles, M.A., & Tsang, E.W.K. (2008). Inter–organizational knowledge transfer: current themes and future prospects. *Journal of Management Studies, 45*, 677–690.

Ellison, G., Glaeser, E.L. (1997). Geographic concentration in U.S. manufacturing industries: A dartboard approach. *Journal of Political Economy, 105*, 889–927.

Eurostat. (2008). *Statistical Classification of Economic Activities in the European Community*. Luxembourg: European Communities.

Feenstra, R.C. (2000). *World Trade Flows 1980–1997*. Working paper, Center for International Data, Institute of Governmental Affairs, University of California, Davis.

Accepted Article

This article is protected by copyright. All rights reserved.

Feenstra, R.C., Lipsey, R.E., Deng, H., Ma, A.C., & Mo, H. (2005). *World Trade Flows: 1962–2000*. NBER Working Paper No. 11040.
Glaeser, E.L. (2010). Introduction. In E.L. Glaeser (Ed.), *Agglomeration Economics* (pp. 1–14). The University of Chicago Press.
Goerzen, A., Asmussen, C.G., & Nielsen, B. (2013). Global cities and multinational enterprise location strategy. *Journal of International Business Studies, 44*, 427–450.
Head, K., Ries, J., & Swenson, D. (1995). Agglomeration benefits and location choice: Evidence from Japanese manufacturing investments in the United States. *Journal of International Economics, 38*, 223–247.
Hennart, J–F., & Park, Y–R. (1994). Location, governance, and strategic determinants of Japanese manufacturing investment in the United States. *Strategic Management Journal, 15*, 419–436
Hernandez, E. (2014). Finding a home away from home: Effects of immigrants on firms' foreign location choice and performance. *Administrative Science Quarterly, 59*, 73–108.
Hernandez, E., Sanders, W.G., & Tuschke, A. (2015). Network defense: Pruning, grafting, and closing to prevent leakage of strategic knowledge to rivals. *Academy of Management Journal, 58*, 1233–1260.
Hervas–Oliver, J–L., Sempere–Ripoll, F., Alvarado, R.R., & Estelles–Miguel, S. (2018). Agglomeration and firm performance: Who benefits and how much? *Regional Studies, 52*, 338–349.
Hoetker, G. (2007). The use of logit and probit models in strategic management research: critical issues. *Strategic Management Journal, 28*, 331–343.
ISTAT, ICE. (2002). *Commercio estero e attività internazionali delle imprese. Annuario 2002. L'Italia nell'economia internazionale*. Roma: ICE.
Johanson, J., & Vahlne, J–E. (2009), The Uppsala internationalization process model revisited: From liability of foreignness to liability of outsidership. *Journal of International Business Studies, 40*, 1411–1431.
Jost, P.J., & Velden, C. (2008). Organizational design of R&D after mergers and the role of budget responsibility. *Journal of Economics and Business, 60*, 469–484.
Kenney, M., & Florida, R. (1992). The Japanese transplants: Production organization and regional development. *Journal of the American Planning Association, 58*, 20–38.
Koçak, Ö., & Özcan, S. (2013). How does rivals' presence affect firms' decision to enter new markets? Economic and sociological explanations. *Management Science, 59*, 2586–2603.
Li, P., & Bathelt, H. (2018). Location strategy in cluster networks. *Journal of International Business Studies, 49*, 967–989.
Mariotti, S., Mutinelli, M., & Sansoucy, L. (2014). *Italia multinazionale 2014. Le partecipazioni italiane all'estero ed estere in Italia*. Soveria Mannelli: Rubbettino Editore.
Mariotti, S., & Piscitello, L. (1995). Information costs and location of FDIs within the host country: Empirical evidence from Italy. *Journal of International Business Studies, 26*, 815–841.
Mariotti, S., Piscitello, L., & Elia, S. (2010). Spatial agglomeration of multinational enterprises: The role of information externalities and knowledge spillovers. *Journal of Economic Geography, 10*, 519–538.
Marshall, A. (1890). *Principles of Economics*. London: MacMillan.
McCann, B., & Folta, T. (2008). Location matters: Where we have been and where we might go in agglomeration research. *Journal of Management, 34*, 532–565.

Accepted Article

This article is protected by copyright. All rights reserved.

McCann, B., & Folta, T. (2011). Performance differentials within geographic clusters. *Journal of Business Venturing, 26*, 104–123.
McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). New York: Academic Press.
Neffke, F., Svensson, H., Boschma, R.A., Lundquist. K–J., & Olander, L–O. (2011). The dynamics of agglomeration externalities along the life cycle of industries. *Regional Studies, 45*, 49–65.
Potter, A., & Watts, H.D. (2011). Evolutionary agglomeration theory: Increasing returns, diminishing returns, and the industry life cycle. *Journal of Economic Geography, 11*, 417–55.
Puga, D. (2010). The magnitude and causes of agglomeration economies. *Journal of Regional Science, 50*, 203–219.
Richardson, H.W. (1995). Economies and diseconomies of agglomeration. In H. Giersch (Ed.), *Urban Agglomeration and Economic Growth* (pp. 123–155). Berlin: Springer.
Rigby, D.L., & Brown, W.M. (2015). Who benefits from agglomeration? *Regional Studies, 49*, 28–43.
Rosenthal, S.S., & Strange, W.C. (2001). The determinants of agglomeration. *Journal of Urban Economics, 50*, 191–229.
Rosenthal, S.S., & Strange, W.C. (2004). Evidence on the nature and sources of agglomeration economies. In J.V. Henderson & J–F. Thisse (Eds.), *Handbook of Urban and Regional Economics* (pp. 2119–2171). Amsterdam: Elsevier.
Shaver, J.M., & Flyer, F. (2000). Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic Management Journal, 21*, 1175–1193.
Shaver, J.M., Mitchell, W., & Yeung, B. (1997). The effect of own–firm and other–firm experience on foreign direct investment survival in the United States, 1987–92. *Strategic Management Journal, 18*, 811–824.
U.S. Department of Commerce. (1988). *Foreign Direct Investment in the United States, 1987 Transactions*. U.S. Department of Commerce, International Trade Administration: Washington, DC.
Wang, L., Madhok, A., & Xiao Li, S. (2014). Agglomeration and clustering over the industry life cycle: Toward a dynamic model of geographic concentration. *Strategic Management Journal, 35*, 995–1012.

Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

Table 1 - Description of dependent and explanatory variables in our study *versus* SF original study

SF variables (US, 1987–1995)		Our variables (Italy, 2001–2009)		
Name	***Description***	***Name***	***Description***	***Differences***
SURVIVAL	Whether or not the plant was still controlled by the investing firm in 1995 (i.e., 8 years after the creation)	SURVIVAL	Whether or not the plant was still controlled by the investing firm in 2009 (i.e., about 8 years after the creation)	No
PROPORTION	Proportion of industry establishment in the target state in 1987	PROPORTION	Proportion of industry establishment in the target region, in 2001	No
INDUSTRY GROWTH	Proportionate change in industry shipments (1987–1993)	INDUSTRY GROWTH	Proportionate change in industry shipments (2001–2007)	No
C4	Four–firm concentration ratio, 1987	C4	Four–firm concentration ratio based on employees, 2001	Minor
RELATED	Whether or not the foreign parent controls operations in the same two–digit SIC in the US, in 1986	RELATED	Whether or not the foreign parent controls operations in the same two–digit Nace sector in Italy in 2000	No
US–FOREIGN	Proportion of US firms in the industry with international operations, 1987	IT–FOREIGN	Proportion of Italian firms in the industry with international operations, 2001	No
FINANCIAL DATA	Are financial data available for the foreign parent, in 1986	FINANCIAL DATA	Are financial data available for the foreign parent, in 2000	No
SUBSIDIARIES	Number of subsidiaries (including manufacturing, sales and distribution) controlled by the foreign parent, in the US, in 1986	SUBSIDIARIES	Number of subsidiaries (including manufacturing, sales and distribution) controlled by the foreign parent in Italy, in 2000	No
GEOGRAPHIC CONCENTRATION	Ellison and Glaeser's (1997) 'gamma' for 1987.	GEOGRAPHIC CONCENTRATION	Ellison and Glaeser's (1997) 'gamma' for 2001	No
LARGE ESTABLISHMENT	Indicates whether the new plant's value is greater than average establishment capital stock within the industry in 1987	LARGE ESTABLISHMENT	Indicates whether the new plant's number of employees is greater than average number of employees within the industry in 2001	Minor
PARENT'S NATIONALITY DUMMIES	Foreign parent is Canadian, French, German, Japanese, from the UK	PARENT'S NATIONALITY DUMMIES	Foreign parent is Chinese, French, German, Japanese, Swiss, from the UK, from the US	No
		EXCELLENCE	Whether or not Balassa (1965) Revealed Comparative Advantage (RCA) index in the	New

This article is protected by copyright. All rights reserved.

Accepted Article

			focal sector the year before the entry is larger than 1.5.	
		YEAR	Entry year (2000=-1, 2001=0, 2002=1)	New

This article is protected by copyright. All rights reserved.

Accepted Article

Table 2 - Correlation matrix and other descriptive statistics (n = 113)

	VARIABLE	1	2	3	4	5	6	7	8	9	10	11	12
1	*SURVIVAL*	1											
2	*PROPORTION*	0.22	1										
3	*INDUSTRY_GROWTH*	0.01	0.08	1									
4	*C4*	0.09	-0.01	-0.03	1								
5	*RELATED*	-0.03	-0.16	0.09	0.05	1							
6	*IT-FOREIGN*	0.10	0.07	-0.17	0.09	0.23	1						
7	*FINANCIAL DATA*	-0.18	-0.15	-0.24	0.10	0.12	0.12	1					
8	*SUBSIDIARIES*	0.01	-0.15	0.05	-0.01	0.48	-0.08	0.12	1				
9	*GEOGRAPHIC_CONCENTRATION*	0.02	-0.11	-0.03	0.36	0.13	-0.04	0.07	0.10	1			
10	*LARGE_ESTABLISHMENT*	-0.21	-0.08	0.04	-0.32	0.13	-0.20	0.15	0.18	-0.14	1		
11	*EXCELLENCE*	-0.24	-0.05	-0.11	-0.07	-0.02	0.18	-0.01	-0.05	0.08	-0.09	1	
12	*YEAR*	0.11	0.03	0.09	-0.06	0.15	-0.03	-0.02	0.12	-0.02	-0.11	-0.10	1
	Mean	0.578	0.155	0.168	13.13	0.222	0.578	0.956	2.976	0.029	0.333	0.467	-0.078
	Standard Deviation	0.497	0.116	0.267	12.32	0.418	0.123	0.207	4.918	0.030	0.474	0.502	0.810
	Minimum	0	0.003	-0.618	0.55	0	0.308	0	0	0.000	0	0	-1
	Maximum	1	0.401	0.793	76.81	1	0.760	1	19	0.177	1	1	1

This article is protected by copyright. All rights reserved.

Table 3 - Conditional logit results of firms' location choice

	1. Choice set: 20 regions	2. Choice set: 20 regions; Geographically concentrated industries	3. Restricted choice set: 15 regions
Agglomeration measure			
PROPORTION	9.151	8.791	3.554
	(10.97)	(7.59)	(1.91)
Other variables			
Fixed effects	Not included	Not included	Included
Number of cases	90	43	90
Number of alternatives	20	20	15
Wald χ_1^2 (p–value)	120.40 (0.000)	57.62 (0.000)	3.65 (0.056)
Log likelihood	-218.54	-103.56	-193.45

t-ratios in parentheses.

Table 3b - Conditional logit results of firms' location choice (logarithmic version)

	1. Choice set: 20 regions	2. Choice set: 20 regions; Geographically concentrated industries	3. Restricted choice set: 15 regions
Agglomeration measure			
PROPORTION	0.964	0.880	0.594
	(9.55)	(6.48)	(2.49)
Other variables			
Fixed effects	Not included	Not included	Included
Number of cases	90	43	90
Number of alternatives	20	20	15
Wald χ_1^2 (p–value)	91.17 (0.000)	41.96 (0.000)	6.20 (0.013)
Log likelihood	-207.88	-100.74	-192.04

t–ratios in parentheses.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 4 - Logistic regression results: dependent variable = *Survival*

	1. Full sample	2. Geographically concentrated industries
Agglomeration measures		
PROPORTION	3.381	2.415
	(1.47)	(0.67)
	[0. 817]	[0. 596]
Other variables		
INTERCEPT	-2.337	-5.186
	(-0.52)	(-1.88)
INDUSTRY GROWTH	0.298	0.032
	(0.32)	(0.03)
	[0.072]	[0.008]
C4	0.019	-0.017
	(0.93)	(-0.67)
	[0.005]	[-0.004]
RELATED	-0.782	-2.252
	(-0.98)	(-1.13)
	[-0.189]	[-0.559]
IT-FOREIGN	1.995	7.138
	(0.88)	(1.59)
	[0.482]	[1.772]
FINANCIAL DATA	n.a.*	n.a.*
	(n.a.)	(n.a.)
	[n.a.]	[n.a.]
SUBSIDIARIES	0.034	0.356
	(0.49)	(2.10)
	[0.008]	[0.088]
GEOGRAPHIC CONCENTRATION	1.773	4.601
	(0.02)	(0.39)
	[0.042]	[1.142]
YEAR	0.305	-0.399
	(0.98)	(-0.74)
	[0.074]	[-0.099]
Country dummies variables included	Yes	Yes
n (survived)	90 (52)	43 (23)
Wald χ^2_{15} (p–value)	11.39 (0.66)	9.68 (0.72)
Log pseudolikelihood	-55.26	-22.69

Accepted Article

This article is protected by copyright. All rights reserved.

t–ratios in parentheses; marginal effects, evaluated at the mean of independent variables, are in square brackets.
* The parameter cannot be estimated since there are only two cases with *FINANCIAL DATA=0*, and in both cases *SURVIVAL=1* (perfect prediction).

Accepted Article

This article is protected by copyright. All rights reserved.

Table 4b - Logistic regression results: dependent variable = *Survival*

	1. Full sample	1b. Full sample	2. Geographically concentrated industries
Agglomeration measures			
PROPORTION	7.194 (1.91) [1.716]	6.565 (1.94) [1.569]	11.722 (1.58) [2.804]
EXCELLENCE	0.401 (0.43) [0.096]		
*PROPORTION*EXCELLENCE*	-7.094 (-1.40) [-1.692]	-5.362 (-1.73) [-1.281]	-11.205 (-1.86) [-2.680]
Other variables			
INTERCEPT	-2.234 (-1.31)	-2.038 (-1.27)	-5.447 (-1.96)
INDUSTRY GROWTH	0.342 (0.36) [0.082]	0.291 (0.30) [0.070]	-0.062 (-0.05) [-0.015]
C4	0.015 (0.69) [0.004]	0.012 (0.57) [0.003]	-0.021 (-0.81) [-0.005]
RELATED	-0.627 (-0.80) [-0.150]	-0.640 (-0.84) [-0.153]	-2.266 (-1.26) [-0.542]
IT–FOREIGN	2.023 (0.86) [0.483]	1.978 (0.85) [0.473]	8.316 (1.87) [1.989]
FINANCIAL DATA	n.a.* (n.a.) [n.a.]	n.a.* (n.a.) [n.a.]	n.a.* (n.a.) [n.a.]
SUBSIDIARIES	0.027 (0.38) [0.006]	0.027 (0.40) [0.007]	0.381 (2.30) [0.091]
GEOGRAPHIC CONCENTRATION	0.531 (0.05) [1.127]	0.714 (0.07) [1.171]	0.550 (0.05) [0.131]
YEAR	0.265 (0.81) [0.063]	0.287 (0.88) [0.069]	-0.181 (-0.33) [-0.043]
Country dummies	Yes	Yes	Yes
n (survived)	90 (52)	90 (52)	43 (23)
Wald χ^2 (p–value)	12.82 (0. 69)	13.25 (0.58)	11.48 (0.65)
Log pseudolikelihood	-53.42	-53.51	-20.87

Accepted Article

This article is protected by copyright. All rights reserved.

t-ratios in parentheses; marginal effects, evaluated at the mean of independent variables, are in square brackets. Marginal effects for *PROPORTION* and *EXCELLENCE* not computed due to the interaction term.
* The parameter cannot be estimated since there are only four cases with *FINANCIAL DATA=0*, and in all cases *SURVIVAL=1* (perfect prediction).

Table 5 - Conditional logit results of foreign firms' location choice (LARGE_ESTABLISHMENT as a proxy of the strength of the entrant MNE)

	1. Choice set: 20 regions	2. Choice set: 20 regions; Geographically concentrated industries	3. Restricted choice set: 15 regions
Agglomeration measure			
ln(PROPORTION)	0.959	0.862	0.596
	(7.84)	(5.81)	(2.44)
ln(PROPORTION) ×	0.017	0. 110	-0.006
LARGE_ESTABLISHMENT	(0.08)	(0.30)	(-0.03)
Other variables			
Fixed effects	Not included	Not included	Included
Number of cases	90	43	90
Number of alternatives	20	20	15
Wald χ_2^2 (p–value)	91.20 (0.000)	42.05 (0.000)	6.20 (0.045)
Log likelihood	-207.87	-100.69	-192.04

t–ratios in parentheses.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 5b - Conditional logit results of foreign firms' location choice and heterogeneity of local context (LARGE_ESTABLISHMENT as a proxy of the strength of the entrant MNE)

	1. Choice set: 20 regions	2. Choice set: 20 regions; Geographically concentrated industries	3. Restricted choice set: 15 regions
Agglomeration measure			
ln(PROPORTION)	0.748	0.453	0.287
	(5.08)	(2.60)	(1.01)
ln(PROPORTION) × *LARGE_ESTABLISHMENT*	0.381	0.700	0.301
	(1.25)	(1.40)	(1.06)
ln(PROPORTION) × *EXCELLENCE*	0.649	1.572	0.638
	(2.17)	(2.96)	(2.14)
ln(PROPORTION) × *LARGE_ ESTABLISHMENT* × *EXCELLENCE*	-0.938	-1.982	-0.822
	(-2.02)	(-2.33)	(-1.86)
Other variables			
Fixed effects	Not included	Not included	Included
Number of cases	90	43	90
Number of alternatives	20	20	15
Wald χ^2 (p–value)	85.31 (0.000)	31.61 (0.000)	9.47 (0.050)
Log likelihood	-204.81	-93.09	-189.02

t–ratios in parentheses.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 5c - Conditional logit results of foreign firms' location choice and heterogeneity of local context (EXCEL_COMP as a proxy of the strength of the entrant MNE)

	1. Choice set: 20 regions	2. Choice set: 20 regions; Geographically concentrated industries	3. Restricted choice set: 15 regions
Agglomeration measure			
ln(PROPORTION)	0.735	0.345	0.230
	(5.48)	(2.00)	(0.80)
ln(PROPORTION) × *EXCEL_COMP*	1.176	1.479	1.128
	(2.01)	(2.26)	(1.88)
ln(PROPORTION) × *EXCELLENCE*	0.447	1.312	0.468
	(1.77)	(2.72)	(1.88)
ln(PROPORTION) × *EXCEL_COMP* × *EXCELLENCE*	-1.299	-1.652	-1.248
	(-1.89)	(-1.74)	(-1.80)
Other variables			
Fixed effects	Not included	Not included	Included
Number of cases	90	43	90
Number of alternatives	20	20	15
Wald χ^2 (p–value)	85.09 (0.000)	33.89 (0.000)	11.16 (0.025)
Log likelihood	-203.86	-91.67	-188.02

t–ratios in parentheses.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 5d - Conditional logit results of foreign firms' location choice and heterogeneity of local context (LARGEorEXCEL as a proxy of the strength of the entrant MNE)

	1. Choice set: 20 regions	2. Choice set: 20 regions; Geographically concentrated industries	3. Restricted choice set: 15 regions
Agglomeration measure			
ln(PROPORTION)	0.648	0.314	0.165
	(4.22)	(1.70)	(0.56)
ln(PROPORTION) × *LARGEorEXCEL*	0.584	0.935	0.500
	(2.05)	(2.24)	(1.82)
ln(PROPORTION) × *EXCELLENCE*	1.018	1.698	1.051
	(2.43)	(2.66)	(2.40)
ln(PROPORTION) × *LargeorEXCEL* × *EXCELLENCE*	-1.325	-1.662	-1.264
	(-2.55)	(-1.97)	(-2.40)
Other variables			
Fixed effects	Not included	Not included	Included
Number of cases	90	43	90
Number of alternatives	20	20	15
Wald χ^2 (p–value)	85.45 (0.000)	34.96 (0.000)	11.64 (0.020)
Log likelihood	-202.94	-92.36	-187.12

t–ratios in parentheses.

Accepted Article

This article is protected by copyright. All rights reserved.

Table 6 - Distribution of EXCELLENCE and EXCEL_COMP in the 90 investments included in our dataset (greenfield investments in Italy 2001)

	EXCELLENCE=0	EXCELLENCE=1	ALL
EXCEL_COMP=0	51.1%	28.9%	80.0%
EXCEL_COMP=1	7.8%	12.2%	20.0%
ALL	58.9%	41.1%	100.0%

Table 7 - Distribution of EXCELLENCE and EXCEL_COMP in the 121 investments related to SF dataset (greenfield investments in US 1987, RCA based on 1996 data)

	EXCELLENCE=0	EXCELLENCE=1	ALL
EXCEL_COMP=0	44.7%	9.6%	54.4%
EXCEL_COMP=1	32.5%	13.2%	45.6%
ALL	77.2%	22.8%	100.0%

Table 8 - Coefficient of PROPORTION in the location equation as a function of EXCELLENCE and EXCEL_COMP (based on Table 5c, column 1)

	EXCELLENCE=0	EXCELLENCE=1
EXCEL_COMP=0	0.735	1.182
EXCEL_COMP=1	1.911	1.059

Table 9 - Distribution by sector and country of origin of the 37 cases in the US 1987 sample where EXCELLENCE=1 and EXCEL_COMP=0

SECTOR	Number of investments from:				
	Canada	Germany	Japan	UK	France
242 Sawmills and Planing Mills	1				
265 Paperboard Containers and Boxes		1			
285 Paints and Allied Products		1			
345 Screw Machine Products, Bolts, etc.		1			
354 Metalworking Machinery		1	2		
355 Special Industry Machinery		1	2		
359 Industrial Machinery, Nec				1	
362 Electrical Industrial Apparatus			1		
369 Misc. Electrical Equipment & Supplies			3		
371 Motor Vehicles and Equipment	1		15		
373 Ship and Boat Building and Repairing			1		
386 Photographic Equipment and Supplies			4		
395 Pens, Pencils, Office, & Art Supplies			1		
TOTAL	2	5	29	1	0

Accepted Article

This article is protected by copyright. All rights reserved.

Figure 1 - Estimated probabilities of location choice *versus PROPORTION* (Linear model = probabilities based on the model in Table 3, column 1; Logarithmic model = probabilities based on the model in Table 3b, column 1; the plot to the right zooms on the cases where *PROPORTION*<0.1)

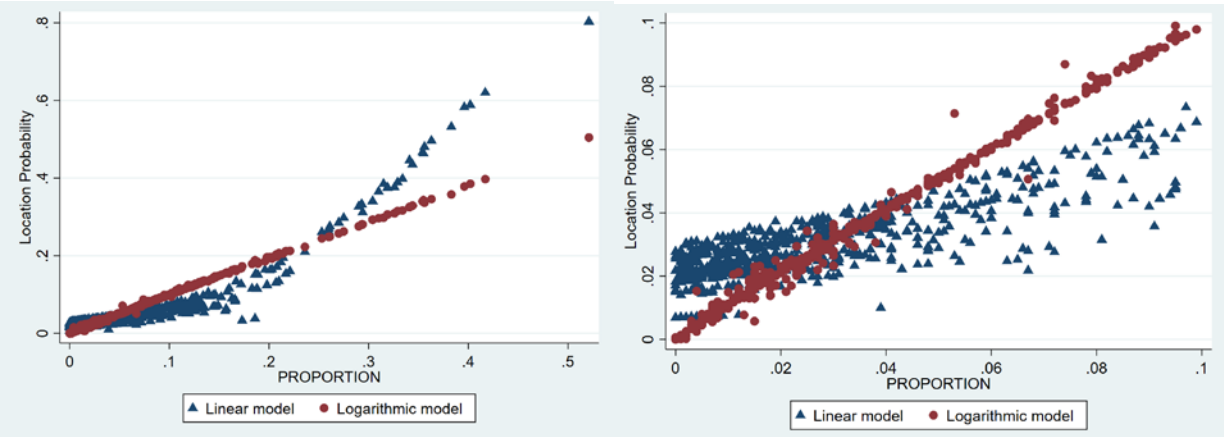
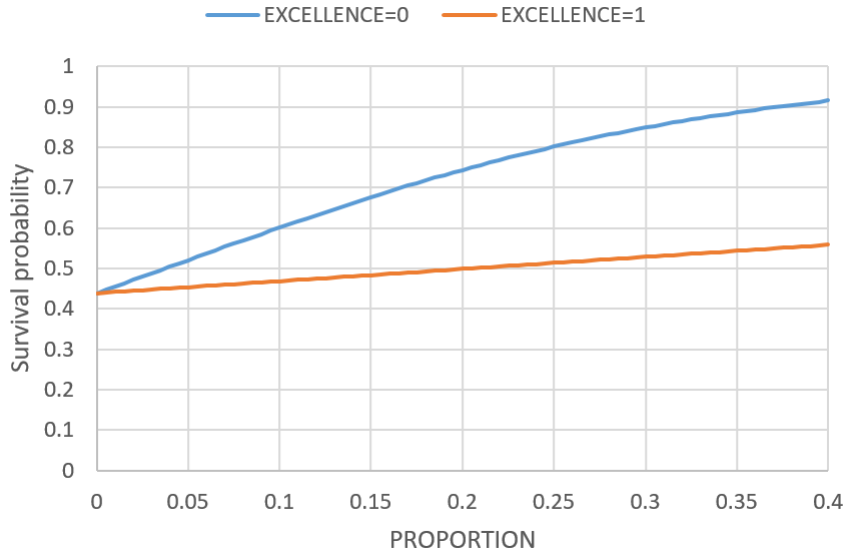


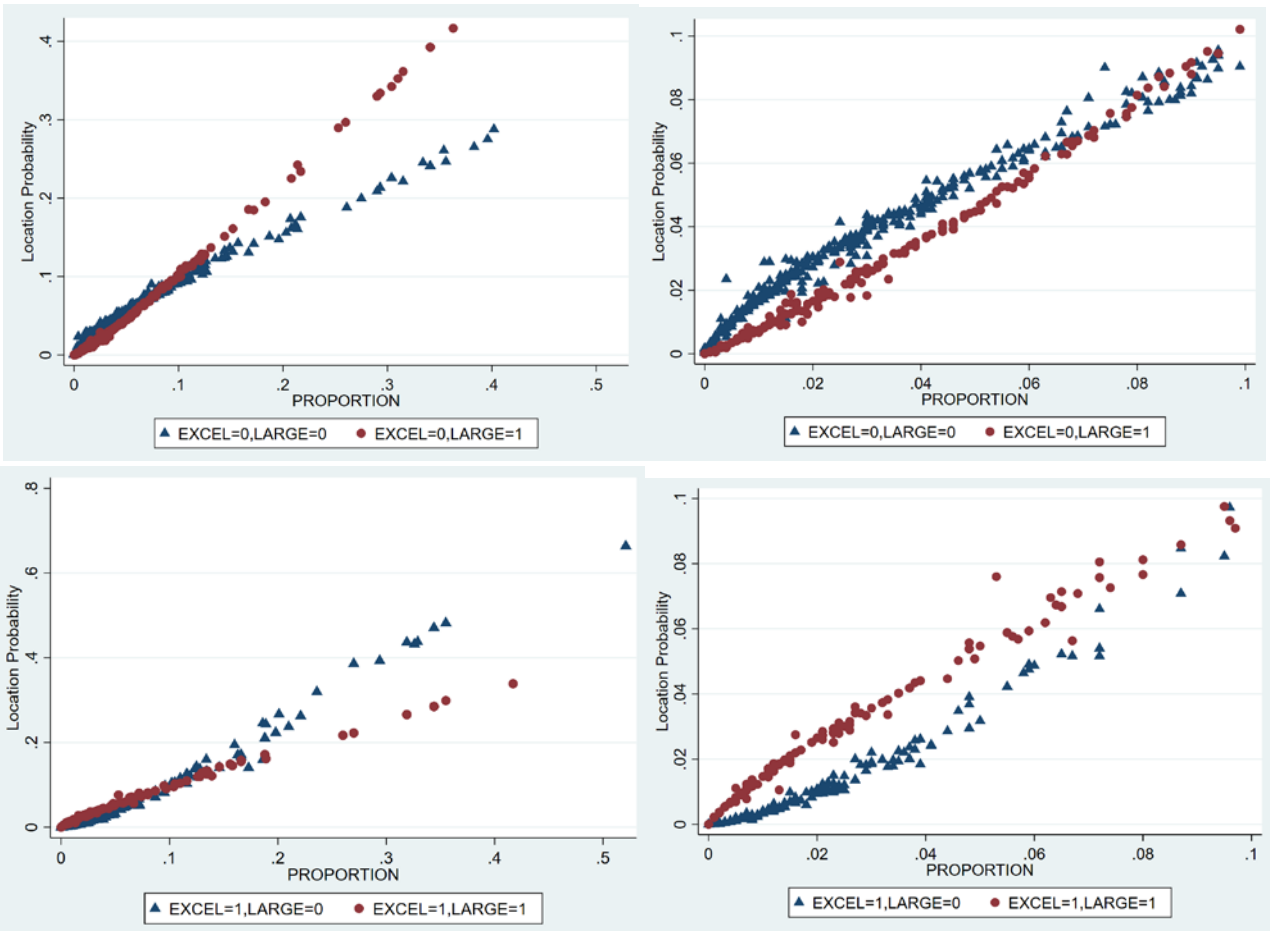
Figure 2 - Survival probability as a function of *PROPORTION* (The curve is based on the model in Table 4b, column 1b; all the other variables in the model apart from *PROPORTION* have been fixed to represent a typical situation, i.e., to the mode for the dummies and to the median for the other variables; the country is set to US)



Accepted Article

This article is protected by copyright. All rights reserved.

Figure 3 - Estimated probabilities of location choice *versus PROPORTION* (based on the model in Table 5b, column 1b; the plots to the right zoom on the cases where *PROPORTION*<0.1)



Accepted Article

This article is protected by copyright. All rights reserved.

Accepted Article

Appendix

Table 1A - Characteristics of Italian regions, 2001

	A: Land area (,000 km2)	B: Population (,000)	C: Number of Establishments - manufacturing (,000)	D: Establishments density (C/A)
Mean	15.10	3033.3	29.52	1.74
Median	14.45	1814.3	18.44	1.59
St.Deviation	7.26	2468.7	28.98	1.20
Skewness	-0.01	1.03	1.74	1.14
Kurtosis	-1.37	1.32	4.45	1.99
Min	3.26	127.3	1.01	0.31
Max	25.83	10008.3	124.39	5.21
Correlation with B	0.70			
Correlation with C	0.64	0.90		
Correlation with D	0.37	0.80	0.91	

Source: ISTAT, Italian Industry and Services Census.

This article is protected by copyright. All rights reserved.