Agglomerations and firm performance: who benefits and how much?

1. Introduction

How does a firm's knowledge moderate access to localization economies? Who gains and who loses from co-location? The conclusions from the debate are far from clear thanks to an unresolved tension stemming from the consideration of firms' internal knowledge in the localization economies topic: while more knowledge-intensive firms have more to lose and less to gain (Rigby and Brown, 2013; Shaver and Flyer, 2000), it is also said that firms with higher knowledge stocks benefit more from agglomeration (McCann and Folta, 2011). The present study responds to the recent calls of Rigby and Brown (2013), or MCCann and Folta (2011), among others, to carry on exploring how co-located firms benefit from agglomeration.

Moving down from the regional level to the firm level as a unit of analysis, this paper focuses on how heterogeneous firms, with different routines, capabilities, resources and knowledge, that is, firm-specific assets or capabilities that are internal, capture externalities and, if so, whether these externalities are or are not asymmetrically gained within agglomerations. This work contributes to the geography of innovation strands and advances the research on localization economies, showing how firms with different innovation capabilities achieve localization gains differently. In other words, this study sheds light on the topic of whether returns within agglomerations and their localization externalities are asymmetrically and unevenly distributed among co-located firms.

The rationale of this paper is as follows. Despite substantial amounts of work on the relationship between localization externalities and firm performance (e.g. Sorenson and Audia, 2000; McCann and Folta, 2011; Belussi and Hervas-Oliver, 2017), important issues remain unresolved and the debate is far from conclusive. There are studies which have found that localization has no effect or even has negative effects on performance (e.g. Gilbert et al., 2008; Audia and Rider, 2010), while others have found the link to be positive (e.g Folta, Cooper and Baik, 2006; Hervas-Oliver and Albors-Garrigos, 2009; McCann and Folta, 2011). Moreover, another unresolved discussion has focussed on the potential asymmetric benefits for located firms. While it is agreed that not all firms benefit equally from being located in an agglomeration (e.g. Baum and Haveman, 1997; Rigby and Brown, 2013), some studies have concluded that strong firm-specific assets or knowledge-rich firms are the main beneficiaries (e.g. McCann and Folta, 2011), while others say that, on the contrary, it is weaker firm-specific assets or knowledge-poor firms which gain the most (Shaver and Flyer, 2000; Rigby and Brown, 2013). Again, there is no conclusive evidence. As such, our study attempts to shed light on these two different yet intertwined debates.

Our point of departure is localization externalities[1] occurring in agglomerations, and our overall aim is to evaluate the role of those externalities on a firm's innovative performance. In particular, this paper seeks to provide empirical evidence on (i) whether co-location in an agglomeration improves innovation, that is, whether agglomeration benefits exist and, (ii) if so, how are they distributed among agglomerated firms, that is, equally or asymmetrically.

Our study advances economic geography by intersecting it with the strategy literature and thus untangling the impact of localization externalities on a firm's innovativeness, showing the mechanisms that link agglomeration and innovation and exploring potential asymmetric benefits in agglomerations and the differing groups of firms benefiting from them. The data used in the study encompass an entire country and multiple industries, allowing generalization of results. Using the Community Innovation Survey (CIS), a rich full-scale dataset covering 6,697 Spanish firms for the period 2004-2006 and regional-level data, our predictions are supported. Spain was selected for study because it presents a high proportion of localization economies in Europe[2]. The paper is organized as follows. The following section formalizes our hypotheses. Then, in the third, we elaborate on our data and our empirical design. In the fourth section, the results are presented, together with a discussion. Finally, conclusions are developed and some areas for future research are discussed.

2. Agglomeration and Firm Performance: hypothesis development

The spatial concentration or agglomeration of economic activity leads to the emergence of externalities in many different forms. From the economics perspective, Marshall, Arrow and Romer (MAR) externalities (e.g. Glaeser *et al.*, 1992), focused on firms in

[1] Marshall (1890), Arrow (1962), and Romer (1986) put forward a concept, which was later formalized by the seminal work of Glaeser et al. (1992) and became known as the Marshall–Arrow–Romer (MAR) mode

[2] See Boix, 2009

the same industry, on the one hand, are economies of scale external to the firm but internal to a territorial system, facilitating agglomeration, due to the advantages of localization, such as the reduction of production costs, access to specialized inputs and suppliers and the better access to learning due to the presence of knowledge or technology spillovers (e.g. Audretsch and Feldman, 1996). The industrial district literature (MID), on the other hand, points out the *social dimension* (e.g., Becattini, 1990; Brusco, 1982) of agglomerations, facilitating a model of network-based and flexible specialization of production fueled by subcontracting, fostering the co-existence of competition and co-operation that positively impacts innovation (e.g. Camagni, 1991; Hervas-Oliver and et al., 2017). Overall, and considering both perspectives, *localization externalities* allow geographically concentrated firms in the same industry to learn from one other, to exchange ideas, to employ imitation business interactions, as well as accessing external knowledge and resources without monetary transactions (Saxenian, 1994), that is to say, unintentional and uncompensated exchange of knowledge among firms, thereby helping to configure a firm's specific capabilities and returns (Marshall, 1890:32; Saxenian, 1994; Hervas-Oliver and Albors-Garrigos, 2007). Claims of these effects have been supported by empirical observations of improved firm performance, in respect of innovativeness (Bell, 2005), financial performance (Kukalis, 2010) or patent activity (McCann and Folta, 2011). Localization externalities are the focus of our study.

Agglomerations provide different types of externalities, like the existence of suppliers, abundant skilled labor or knowledge-abundant contexts, among many other factors, favoring inter-firm knowledge exchange (e.g. Tallman et al., 2004; Stuart and Sorenson, 2003). Those externalities and knowledge rich contexts, therefore, might activate the development of absorptive capacity because firms need to build such a capacity in order to profit from their rich environment. As Cohen and Levinthal's (1990) seminal contribution points out, the learning environments in which firms operate condition the propensity to invest in absorptive capacity in order to capture externalities available there. Thus they say: "g*reater technological opportunity signifies greater amounts of external information, which increase the firm's incentive to build absorptive capacity" (1990:142).* Thus, firms located in agglomerations are more likely than outsiders to increase their firm-specific assets or internal knowledge in order to tap into cluster resources. In other words, assuming that agglomerations produce externalities, *ceteris paribus*, a firm co-located in a region characterized by a relatively high level of its industry specialization, is more likely to increase investment in its own firm-specific assets or knowledge to access externalities. In this vein, Arikan (2009:671-673) points out the fact that firms in agglomertions have to develop absorptive capacity to maximize the potential gains that clusters offer. Similarly, Tallman and Chacar (2011) argue that localization externalities[3], which are deeply embedded, tacit, and inherently immobile (sticking to the territory), increase absorptive capacities for the local firms, because firms recognize the potentially available knowledge. The term 'absorptive capacity' not only describes a firm's ability to scan[4] or evaluate information from its environment, but also to the ability to integrate new external knowledge into a firm's internal innovation activities (Cohen and Levintahl, 1990). External knowledge is *realized* and integrated[5] into a firm's repository of knowledge at the point the firm assimilates and applies the new knowledge (Cohen and Levinthal, 1990). The level of absorptive capacity is therefore highly correlated with a firm's *innovation capability,* in the sense of the resource-based view (e.g. Teece et al., 1997). This idea is put forward in Escribano et al., (2009:98) who posit that *the way to isolate the role of absorptive capacity is by studying its moderating effect on the impact of external knowledge flows on innovation performance*, that is, by studying the effects of the complementary combination of internal and external sources of knowledge. Therefore, we posit that the existence of externalities in agglomerations foster the development of firm-specific assets or internal knowledge to access to those potentially available external resources, forming synergies or combinations of internal and external sources of knowledge, assuming that the combination of internal and external knowledge, and the exploitation of synergies between them, facilitates innovation (e.g. Arora and Gambardella, 1990). Therefore, it is expected that internal and external accessed knowledge combination and its impact on a firm's innovative performance is higher for co-located firms, vis-à-vis non-co-located ones, due to the fact that the existence of more external knowledge or resources (skilled labor, common infrastructure, supplies and so forth) will strengthen that effect. On the contrary, mostly because the poor external knowledge and resources

[3] They refer to *communities of practice* in clusters and the firm's organization of resources in order to access that local knowledge.

[4] Scanning capability only refers to assessing external sources of knowledge, as Cassiman and Veugelers (2000) and Arbussa and Coenders (2007) state.

[5] In the sense of Zahra & George (2002).

available, non-co-located firms will present a weaker combination of internal and external knowledge, vis-à-vis co-located firms. Hence, it is expected that internal and external accessed knowledge combination and its impact on a firm's innovative performance will be lower for co-located firms. In econometric terms, we expect a more robust and higher interaction effect (combination between internal and external sources of knowledge) on innovation for co-located firms. All in all we posit that agglomerations provide, therefore, the perfect setting for fostering a complementary combination of internal and external sources of knowledge and, for this reason, we expect that this combination will be reinforced in agglomerations. We use the term 'co-location' for those firms in relatively high own-industry employment regions. Thus, our first hypothesis can be stated as follows, signaling an interaction effect on a firm's innovation:

Hypothesis 1: Agglomerations exert a positive impact of a co-located firm's internal and external knowledge combination on its innovative performance

Subsequently, this study restricts and focuses analysis only on co-located firms, that is, those firms co-located in a relatively high own-industry employment region (those with LQ>1 or just agglomerations in this study), and considers the heterogeneity amongst firms. Firms have the potential to exploit resources, but vary in the extent to which they do so, especially in agglomerations (Hervas-Oliver and Albors-Garrigos, 2009), due to the fact that even though there are the same external resources for all firms, accessing those resources is moderated by a firm's innovation capacity. Asymmetry in firm-specific assets or learning capabilities leads to an asymmetric access of external resources, thus corresponding to an asymmetric distribution of the benefits gained from accessing knowledge in an agglomeration: the same agglomeration in the same industry and region renders different gains to each co-located firm, depending on each firm's knowledge base. Our paper posits that due to existing knowledge, heterogeneity or different innovation capacity is based on a firm-specific internal knowledge amongst firms, and thus, having differing capabilities to access and integrate external knowledge, agglomerated firms do not gain equally from their rich environments. Recent evidence supports this statement (McCann and Folta, 2011), indicating that those firms co-located in agglomerations which have a stronger knowledge base are better able to access and integrate external knowledge, thus gaining the most. This suggests a linear and positive effect between a firm's innovation capacity and its gains in agglomerations. This argument, however, is only partial.

The linearity of the strength of the above argument is clear, but it neglects the fact that knowledge-rich firms can also contribute to agglomerations through knowledge spillovers (e.g. Canina et al., 2005) due to the intense competition existent in agglomerations (Sorenson and Audia, 2000). The argument has been put forward that firms possessing superior technologies and knowledge have an incentive to avoid agglomerations because they have much less to gain (Shaver and Flyer, 2000). In addition to this *less to gain* idea, firms in an agglomeration face increased direct *competition* (Baum and Mezias, 1992) and knowledge flows easily among located firms (e.g. Tallman et al., 2004), imitation is pervasive, and inter-firm worker mobility is a reality (e.g. Saxenian, 1990). Therefore, the competitive dynamics of firms in agglomerations need to encompass also firms' potential *involuntary spillovers* and the effect from a high level of competition. Furthermore, the gain from externalities should be measured as the *net effect* of the potential benefits accrued from learning in agglomerations, on the one hand, and the potential negative costs, on the other hand. In fact, the net effect can even be negative (Baum and Mezias, 1992; Sorenson and Audia, 2000), depending on each firm's assets or knowledge base. The fact that knowledge-rich firms have less to gain from agglomerations (Shaver and Flyer, 2000), and that knowledge-poor firms have more to gain, means that knowledge-rich firms might experience negative net effects from location in agglomerated areas. In all, we expect that, in agglomerations, the net effect of accessing to external sources of knowledge on a firm's innovative performance from localization economies is moderated by a firm's innovative capabilities. Taking the above mentioned opposed arguments, therefore, we posit that there will be asymmetric gains for agglomerated firms. In fact, as we have argued that there are two opposite arguments, one positive and other negative, we suggest that a curvilinear effect exists. Put another way, the interaction effect from the first hypothesis (combination between internal and external knowledge on innovative performance) is unevenly distributed among co-located firms, being dependent on a firm's innovation capability. Therefore, that interaction effect on innovation will be curvilinear and moderated by a firm's innovation capability. Thus, the second hypothesis is developed as follows:

Hypothesis 2: The effect of internal and external knowledge combination on a co-located firm's innovative performance is curvilinear (inverted U-shape), being moderated by a firm's innovation capability

3. Empirical Design

This study utilizes firm-level and regional variables from two different databases. The firm-level data comes from the Spanish CIS 2006, conducted in 2007 and covering the 2004-2006 period. Our empirical analysis covers the effects of introducing innovative activities by *innovatively active* firms (6,697 firms). We control the selection by using Heckman (1979)[6] and also test whether data suffer from common method bias using Harman's single factor test (Greene & Organ, 1973); that is, loading all variables into an exploratory factor analysis and examining the rotated factor solution. No common method variance is identified.

The variables we have used for our analysis are the following: *Inno_product* is a dependent variable which indicates whether an enterprise has introduced a new or improved product or service during the research period. This variable is measured as a dummy variable and has a value of 1 if the firm has introduced a new or improved product and/or service during the studied period, and 0 otherwise. Another variable is *Internal_Capabilities*, which refers to a firm's internal resources of knowledge. The latter is the knowledge base or innovation capability. In constructing this variable we have drawn on the work of Escribano et al., (2009), and also Lane et al. (2006) who emphasize the importance of human resources. Our *Internal_Capabilities* variable is composed from a factor analysis that includes R&D internal expenditures, and the percentage of human resources devoted to R&D in relation to total employees. The resulting scores from a principal component analysis (PCA) represent the absorptive capacities. The two metric variables generating one single component from the analysis, through its scores, represents the dependent variable which explains 52.21 % of the variance (KMO = 0.7172, $p<0.01$). As usual in such analyses, we include control variables, such as Size, measured as the total number of employees, *Industry*

[6] Two-step Heckman procedures check for potential selection problems when restricting the sample to innovative active firms. Thus, one inverse Mills ratio (*lambda* variable) is generated and used for controlling coefficients, which are not significant. Available upon request.

classification, measured using 2-digit NACE-93 industry classification as dummies, and the OECD's classification of low-, medium- and high technology intensive industries.

External sources of knowledge are represented in our analysis by variables which measure the importance respondents gave for the innovation process of the use of external information sources (such as suppliers and customers). These measures capture the role of un-traded interdependencies or externalities from related industries within value chains without monetary transactions (Saxenian, 1994). These variables arose from the question: *how important have the following information sources been for the innovation activities of your enterprise?* (Measured *on a four digit scale from 0 to 3).*

In the agglomeration literature emphasis is put on the role of access to tacit knowledge through interactions (e.g. Becattini, 1990). As such, this paper focuses on *untraded interdependencies* (Storper, 1995), including: learning from interactions with *Suppliers* and *Customers,* and/or through *Trade Associations and* participation in *Events.* By focusing on these four knowledge sources we address the external search strategies of firms and/or the external sources of knowledge they accessed. In a similar way to Laursen and Salter (2006) and with the purpose of using a single indicator for external sources of knowledge (due to methodological requirements below explained with the *logit* corrections on interactions) we constructed this variable as follows. Each of the four sources are coded with either 1 when the firm in question reports that it uses the source to a high degree, and 0 in cases where there is only no, low, or medium use. Afterwards, the scores for the use of the four sources are added up so that each firm gets a score of 0 when no knowledge sources are used to a high degree, while the firm gets the value of 4 when all knowledge sources are used to a high degree (Cronbach's alpha coefficient = 0.71). Firms in the CIS questionnaire are geographically placed on a regional basis according to the location of the enterprises' primary research and development facilities, at NUTS 2. We use the latter information in order to connect CIS data with a regional dataset containing localization indicators.

The regional level data comes from the INE (Spanish Statistics Institute), the same governmental body which administers the CIS itself. The specific source is the *2001 Census* of firms, which is presented using NACE-93 industry classification for each region (Spain comprises 17 regions plus Ceuta and Melilla, which are small cities in the

Northern part of Africa not included in the study[7]). The location quotient is defined as LQ = /Lij/Li)/(Lj/L) where Lij is the number of jobs in the industry i in a region j, Li is the total number of jobs in the industry i in the country, Lj is the number of jobs in a region j, and L is the total number jobs in the country. If the LQ is more than 1 the region is more specialized in an industry than the country's average and so we would conclude that that industry benefits from Marshallian localization economies (Bergman and Feser 1999, Porter, 2003). Basing the LQs on 2001 information limits any possible simultaneity bias; whenever possible the regional indicators are measured before the reference period of the CIS data (2004-2006).

In addition, we also include other measures of potential externalities and control variables, such as *Diversity*, controlling for Jacobian economies (Jacobs, 1960) in the regions, measured by the inverse of the Hirschman-Herfindahl index (HHI),

$$HH_i = 1 \Big/ \left[\sum_{j=1}^{J} \left(\frac{F_{i,j}}{F_i} \right)^2 \right]$$

calculated for the number of firms in all industries for each region, using Census data for 2001, where *F* is the number of firms' local units in each industry, *i* refers to the region, and *j* to the sector.

Additionally, we also include *Population* size and *Density* of the region variables, from data sourced from INE with reference to 2001. The study includes a total of 6,697 firms ranged over 181 industries. The LQs are calculated for a matrix of the 181 industries at 3 digit NACE codes for 17 regions. Regional dummies also control for regional heterogeneity.

From figure A-1 in the Appendix, we observe how our methodology is very sensible reflecting the real composition of industry specialization agglomerations in the Spanish regions. As observed in the matrix, we can easily identify common trends from the Spanish agglomeration facts (e.g. See Boix, 2009[8]).

---

[7] The regions are Andalucía 01; Aragón 02; Asturias 03; Balears 04; Canarias 05; Cantabria 06; Castilla y León 07;Castilla-La Mancha 08; Cataluña 09; Comunidad Valenciana 10; Extremadura 11; Galicia 12; Madrid 13; Murcia 14; Navarra 15; País Vasco 16; Rioja 17.

[8] Boix, R. (2009)

4. Results

Table 1 shows the descriptive statistics and correlation matrix. See table 1.

Insert table 1 about here

Cross-section data restrains from unobserved firm heterogeneity. Besides, the reverse causality between innovation and internal and external sources is also a major point to tackle. In our views, the only way we can control for this potential reverse causality is the way in which the questionnaire is designed. In fact, the CIS questionnaire is designed in a way that perfectly distinguished innovation outputs and inputs and, clearly, the variables utilized for capturing internal and external sources (independent variables) are clearly formulated and contextualized in a way that clearly represents inputs or explanatory variables. Similarly, the innovation output or dependent variables are also clearly formulated and contextualized in a way that clearly represents an output variable. Literature on this topic (see McCann and Folta, Rigby and Brown, and many others) has followed a similar formulation, albeit with different objectives, samples and countries. In any case, our study has no available instruments to check for endogeneity in our model formulation. Please, visit information about the CIS questionnaire to check our argument[9].

Insert table 2 here

As our dependent variable (the introduction of product innovation) is binary, we have based our estimations on a *logit* model. In table 2, the general model is observed in six specifications. The first three ones show individual effects and the general model, while in the last three specifications (from 4 to 6) the first and second hypotheses are tested using the interaction effect and the analysis of its curve. In all the first three specifications, for all firms and then for the two groups of high and low industry specialization (cut-off point established at LQ value 1), we observed how the individual effects for *Internal_Capabilities* and *External_Sources* variables are all positive and

[9] http://ec.europa.eu/eurostat/web/microdata/community-innovation-survey

significantly related to the innovation outcome. These individual effects, however, are not part of the hypothesis but the interaction effect. For this reason, in table 2 throughout the 4, 5 and 6 specifications, we split the sample in two subsamples for the purpose of testing interactions (first hypothesis). The reason of this split is because following Ai and Norton (2003), the nonlinear nature of the logit model means that the marginal effect on an interaction effect is not simply the coefficient (and associated odds ratio) of their interaction. We split the sample into two groups (see McCann and Folta, 2011), addressing firms' locations in LQs lower or higher than 1 because econometrically it is necessary due to the third order effect (*LQ, Internal_Capabilities* and *External_Sources* variables) and its difficulty to be treated in logit regressions. Subsequently, we proceed to apply Ai and Norton's (2003) recommendations to examine the interactions. As mentioned above, the regional literature has usually recommended to set the cut-off point at 1 value (e.g. Bergman and Feser 1999), even though this is just a convention. As regards control variables, such as *Population* and *Density*, it is worth mentioning their poor role capturing the regional additional effects. In general, a negative relationship between *Population* and innovation is observed for LQ>1, and a positive one between *Density* and innovation outcome. In the case of LQ<1 regions, there is not any relationship, but for the whole sample we observe the same as the one encountered for LQ>1. This is explained by the fact that the most specialized regions are also the ones with higher density, irrespective of the population.

First of all, addressing hypothesis 1, we proceed by examining the interaction between internal capabilities (*Internal_Capabilities*) and external knowledge source variables (*Internal_Capabilities X External sources* variable) and its effect on innovation performance or output. This interaction effect is what we have considered as the absorptive capacity, following Escribano et al's, (2009) thoughts. In this part of the analysis the positive or the negative sign is the relevant result, showing or not the complementary combination of internal and external resources depending on the LQs. This produces empirical evidence of positive and statistically significant results related to the moderating effect of localization externalities on that combination. In specifications 4 (addressing the whole sample) and 5 (firms in relatively high own-industry employment regions), the results is positive and statistically significant. In specification 6 (firms not located in relatively high own-industry employment regions), the results are not significant. Corrections are presented graphically in figures 1 to 4 in order to interpret results (following the above mentioned Ai and Norton, 2003). For this reason, in table 2 coefficients are used in order to show graphical corrections, instead of odds ratios. For the sake of brevity we focus on the graphic correction effects of the interactions for interpreting results and the first hypothesis in specifications 5 and 6 from table 2. See table 2.

Insert figures 1 to 4

Figure 1 (table 2, referring to LQs>1, specification 5) shows the size effect of the interaction (LQ>1) when a firm is located in a region characterized by (relatively high own-industry employment) localization externalities. Then, figure 2 (table 2, referring to LQs>1, specification 5) shows the statistical significance of that size effect. The same effects are shown in figure 3 and 4 for the case of firms co-located in relatively low own-industry employment regions (LQs=<1) (table 2, specification 6, LQs=<1). A comparison of figures 1 and 2 with figures 3 and 4 illustrates the different effect that agglomeration exerts on a firm's innovative performance, showing that agglomeration externalities are important and that they matter when combining internal and external knowledge to increase returns. Results reflect the fact that for firms located in agglomerations there is a positive and significant interaction effect on innovative performance (see figures 1 and 2), whereas for firms not co-located in agglomerations (low agglomeration coefficients) the effect is not significant (figures 3 and 4). Thus, it can be stated that a firm's co-location in an agglomeration reinforces and strengthens the complementary internal and external knowledge effect on innovation. See table 2 and figures 1 to 4. The first hypothesis is confirmed. This is done due to the fact that the greater abundance of external sources of knowledge, as well as other externalities, reinforces that combination and its (interaction) effect on innovation, benefiting co-located firms. Whereas, non-co-located firms (or those located in low LQs coefficients) cannot provide as good a combination of internal and external sources of knowledge, vis-à-vis co-located ones, due to the fact that there is less quantity and quality of external resources or externalities available in that particular industry-region territory.

When addressing only the LQs >1 firms (those located in agglomerations) and the shape of the curve, rather than just the positive or negative effect of the interaction, as in the first hypothesis, then the second hypothesis will be tested. Moreover, again in table 2, for firms co-located in agglomerations (LQs>1, specification 5), and in figures 1 and 2, hypothesis 2 is also tested. Figure 1 shows for firms co-located in agglomerations a curvilinear (inverted U-shaped) effect (interaction on innovative performance), reflecting asymmetric gains moderated by firms' innovation capabilities or knowledge base. A firm's innovative capabilities, represented by its propensity to innovate (predicted probability y=1 in the figures), moderate asymmetrically the gains from the complementary combination shown in the first hypothesis in Figure 1. Innovation capability, depicted by a firm's predicted probability to innovate ranges from 0 to 1, the latter being the maximum. This is a good indicator of a firm's knowledge base. All in all, figure 1 highlights a key finding: that the strongest interaction effect on innovative performance (*Internal_Capability x External_Sources* on *Innovative Performance* variables) occurs at the lower end of medium predicted levels of probability of being innovative or knowledge based (approximately 0.3 to 0.6), whereas the effect is less clear-cut for higher levels of predicted probability of being an innovator. Put differently, the middle is the right point to be, because firms with weaker innovation capacity (up to 0.3) present a lower effect on innovative performance, probably due to their lack of accessing external knowledge, that is, a low absorptive capacity. Similarly, stronger firms with a higher innovative capacity (from 0.6 on) have less to gain, since they are already in a pattern of high innovation. Those ranging from 0.3 to 0.6, that is, an average innovative capability or knowledge base, have more to gain, because of the increasing return reported.

In short, throughout the curvilinear (inverted U-shaped) graph shown in figure 1, three groups of firms can be identified: those firms gaining fairly well in innovative performance thanks to being located in agglomerations (innovation capability or predictive level of being innovative between 0.1 and 0.3); those which gain the most (between 0.3 and 0.6); and, finally, those still gaining but gaining the least (from 0.6 to 1). The latter group of firms is the one with the highest innovation capability, but present the lesser gains (compared to the average ones). Also shown is a group of very high innovation capability firms which suffer a loss in innovative performance thanks to being located in agglomerations; this can be seen in the negative values or adverse selection reported by Shaver and Flyer (2000), although these are not statistically significant so as to be sustained. Overall, the striking heterogeneity shown in that curvilinear (inverted U-shape) line reflects and confirms the asymmetric gains or unevenly distributed gains referred to in hypothesis 2. Figure 2 also shows that in the majority of cases the interaction effect is positive and significant at the two-sided 5% level, with some exceptions at the very high end of the predicted probability of being an innovator.

What is the exact meaning of those results? This curvilinear interaction, reflected in an inverted U-shaped graph, shows that the low-medium and medium innovative firms, in terms of innovation capacity, gain the most in terms of innovative output by accessing externalities in agglomerations. This means that, *ceteris paribus*, a relatively high agglomeration of own-industry employment (LQs>1) benefits those firms with lesser probabilities to innovate, that is, the "weaker firms" or "average" firm-specific assets or medium knowledge-base intensive firms. They all have more to gain, while the really weak cannot effectively access, and those very strong in terms of capabilities have less to gain, as not all externalities will benefit them, given the fact that they are very strong in innovation capacity.

In addition, as observed in models 2 and 3 in table 2, the effect of internal capabilities is stronger for non-co-located firms, a fact that confirms that non-co-located firms' innovation patterns are mainly based on internal knowledge, due to the absence of good external resources in their locations. On the contrary, co-located firms use internal knowledge as a means of increasing the absorptive capacity and thus exploiting external knowledge highly abundant in their locations. The fact is even re-confirmed in specifications 5 and 6 in table 2. Extra robustness checks show that collocated firms are more innovative than non-collocated ones (at $p<0.05$; Chi-squared 6.46)[10].

Our results are in line with that part of the literature which has reported evidence of positive and significant links between localization and performance (e.g. Decarolis and Deeds, 1999) Besides, and going a step further, our results also suggest that co-located firms experience asymmetric returns as McCann and Folta (2011) or Rigby and Brown (2013) have already predicted. Our results, however, contradicting those of McCann and Folta (2011), and differing slightly from Rigby and Brown, point out a different yet related inverted U-shape, unfolding thus the variety of possibilities available for co-located firms. Besides, we have tested the data without Catalonia and Madrid, and results coincide, due to the fact that we are not measuring urban but industry specialization externalities. These results are available upon request[11].

[10] We really appreciate one excellent reviewer's suggestion on this point.
[11] We acknowledge reviewers for this point

Lastly, we have conducted a RAMSEY test to check linearity of data. The model is improved adding to *Internal_Capabilities* its quadratic form, but the overall results and hypothesis confirmation are just the same, not affecting our interpretation of results. For the sake of brevity results are available upon request[12].

Acknowledgements

We thank the research project **ECO2015-63645-R** *Open Innovation in Clusters*, funded by the Spanish Ministry of Economics (MINECO and FEDER) for the period 2016-2018. We also thank Dr. Rafael Boix for advice and data on regional issues. I am in debt with Dr. Fiorenza Belussi for comments and ideas.

References

Ai, C., Norton, E. C., 2003, "Interaction terms in logit and probit models". *Economics letters* 80 (1): 123-129.

Arikan, Andaç T., 2009, "Interfirm knowledge exchanges and the knowledge creation capability of clusters". *Academy of Management Review* 34 (4): 658-676.

Arora, Ashish, Gambardella, Alfonso, 1990, "Complementarity and external linkages: the strategies of the large firms in biotechnology". *The Journal of Industrial Economics* 38 (4): 361-379.

Audia, P. and Rider, CH. (2010) Close, but not the same: locally headquartered organizations and agglomeration economies in a declining industry, Research Policy, 39, 360-374

Audretsch, D. B., and Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, 630-640.

Barney, J., 1991, "Firm Resources and Sustained Competitive Advantage". *Journal of Management* 17 (1): 99-120.

Baum, Joel AC, Haveman, Heather A., 1997, "Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990". *Administrative Science Quarterly* 42 : 304-338.

Baum, Joel AC, Mezias, Stephen J., 1992, "Localized competition and organizational failure in the Manhattan hotel industry, 1898-1990". *Administrative Science Quarterly* 37 (4): 580-604.

Becattini, G., 1990, "The Marshallian industrial districs as a socio-economic notion". In *Industrial Districs and Inter-firm cooperation in Italy,* Pike G and Becattini G (eds). International Institute for Labour Studies: Geneva; 413-431

[12] We really acknowledge reviewers for this point

Bell, Geoffrey G., 2005, "Clusters, networks, and firm innovativeness". *Strategic Management Journal* 26 (3): 287-295.

Bell, Simon J., Tracey, Paul, Heide, Jan B., 2009, "The organization of regional clusters". *Academy of Management Review* 34 (4): 623-642.

Belussi, F and Hervas-Oliver, JL (2017) (eds) *Unfolding cluster evolution*, Routledge

Bergman, Edward M., Feser, Edward J., 1999, *Industrial and regional clusters: concepts and comparative applications.* Regional Research Institute, West Virginia University: Virginia.

Boix, R., 2009, The empirical evidence of industrial districts in Spain. In *A Handbook of Industrial Districts,* Becattini G, Bellandi M and De Propis L (Eds). Edward Elgar: MA, USA.; 343-359

Brusco, S (1982). The Emilian model: productive decentralisation and social integration, *Cambridge Journal of Economics* 1982, 6, 167-184.

Camagni, R. (1991). Introduction: from the local 'milieu'to innovation through cooperation networks.*Innovation networks: spatial perspectives*, 1-9.

Canina, Linda, Enz, Cathy A., Harrison, Jeffrey S., 2005, "Agglomeration efects and strategic orientations: Evidence from the US lodging industry". *Academy of management journal* 48 (4): 565-581.

Cohen, W., Levinthal, D., 1990, "Absorptive capacity: a new perspective on learning and innovation". *Administrative Science Quarterly* 35 (1): 128-152.

Escribano, A., Fosfuri, A., Tribo, J. A., 2009, "Managing external knowledge flows: The moderating role of absorptive capacity". *Research Policy* 38 (1): 96-105.

Feldman, Maryann P., 1994, "Knowledge complementarity and innovation". *Small Business Economics* 6 (5): 363-372.

Folta, Timothy B., Cooper, Arnold C., Baik, Yoon-suk, 2006, "Geographic cluster size and firm performance". *Journal of Business Venturing* 21 (2): 217-242.

Gilbert, Brett Anitra, McDougall, Patricia P., Audretsch, David B., 2008, "Clusters, knowledge spillovers and new venture performance: An empirical examination". *Journal of Business Venturing* 23 (4): 405-422.

Glaeser, E., Kallal, H., Scheinkman, J., Shleifer, A., 1992. Growth in cities, Journal of Political Economy 100, 1126-1152.

Greene, C. N., & Organ, D. W. (1973). An evaluation of causal models linking the received role with job satisfaction. *Administrative Science Quarterly, 18*, 95-103.

Heckman, J. J., 1979, "Sample Selection Bias as a Specification Error". *Econometrica* 47 (1): 53-161.

Hervas-Oliver, Jose-Luis, Albors-Garrigos, Jose, 2009, "The role of the firm's internal and relational capabilities in clusters: when distance and embeddedness are not enough to explain innovation". *Journal of Economic Geography* 9 (2): 263-283.

Hervas-Oliver, JL, and *Albors-Garrigos*, J. (*2014*) Are technology gatekeepers renewing clusters? Understanding gatekeepers and their dynamics across cluster life cycles, *Entrepreneurship and Regional Development*, 26, 5-6

Hervas-Oliver, JL, Lleo, M and Cervello, R. (2017) The dynamics of cluster entrepreneurship: Knowledge legacy fromparents or agglomeration effects? The case of the Castellon ceramic tile district, *Research Policy*, 46, 73-92

Jacobs, Jane, 1969, *The economy of cities.* Vintage Books: NewYork.

Jaffe, Adam B., Trajtenberg, Manuel, Henderson, Rebecca, 1993, "Geographic localization of knowledge spillovers as evidenced by patent citations". *Quarterly Journal of Economics* 108 (3): 577-598.

Kukalis, Sal, 2010, "Agglomeration economies and firm performance: the case of industry clusters". *Journal of Management* 36 : 453-481.

Laursen, K., Salter, A., 2006, "Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms". *Strategic Management Journal* 27 (2): 131-150.

Laursen, K., Masciarelli, F., Prencipe, A. (2012), Regions Matter: How Localized Social Capital Affects Innovation and External Knowledge Acquisition, *Organization Science*, Vol. 23, No. 1, January–February 2012, pp. 177–193

Marshall, Alfred, 1920, *Principles of economics: an introductory volume.* Macmillan: London.

Maskell, Peter, 2001, "Towards a knowledge-based theory of the geographical cluster". *Industrial and corporate change* 10 (4): 921-943.

McCann, Brian T., Folta, Timothy B., 2011, "Performance differentials within geographic clusters". *Journal of Business Venturing* 26 (1): 104-123.

McCann, Brian T., Folta, Timothy B., 2008, "Location matters: where we have been and where we might go in agglomeration research". *Journal of Management* 34 (3): 532-565.

McEvily, Bill, Marcus, Alfred, 2005, "Embedded ties and the acquisition of competitive capabilities". *Strategic Management Journal* 26 (11): 1033-1055.

McEvily, B., Zaheer, A., 1999, "Bridging ties: A source of firm heterogeneity in competitive capabilities". *Strategic Management Journal* 20 (12): 1133-1156.

Owen-Smith, J., Powell, W.W., 2004. Knowledge networks as channels and conduits: the effects of spillovers in the Boston biotechnology locale. Organization Science
15, 5–21.

Porter, Michael, 2003, "The economic performance of regions". *Regional Studies* 37 (6-7): 545-546.

Powell, W. W., Koput, K. W., Smith-Doerr, L., 1996, "Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology". *Administrative Science Quarterly* 41 (1): 116-145.

Rigby, D. L., Brown, W. M., 2013, "Who benefits from agglomeration?". *Regional Studies* DOI: 10.1080/00343404.2012.753141 : 1-16.

Romer, P., 1986. Increasing returns and long-run growth. Journal of Political Economy 94, 1002–1037.

Saxenian, AnnaLee, 1994, *Regional advantage: Culture and Competition in Silicon Valley and Route 128.* Harvard University Press: Cambridge, MA.

Saxenian, AnnaLee, 1990, "Regional networks and the resurgence of Silicon Valley". *California Management Review* 33 (1): 89-113.

Shaver, J. Myles, Flyer, Fredrick, 2000, "Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States". *Strategic Management Journal* 21 (12): 1175-1194.

Sorenson, Olav, Audia, Pino G., 2000, "The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States, 1940–19891". *American Journal of Sociology* 106 (2): 424-462.

Storper, M (1995) “The Resurgence of Regional Economies, Ten Years Later: the region as a nexus of untraded interdependencies” European Urban and Regional Studies, Vol. 2 (3).

Stuart, Toby, Sorenson, Olav, 2003, "The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms". *Research policy* 32 (2): 229-253.

Tallman, Stephen, Jenkins, Mark, Henry, Nick, Pinch, Steven, 2004, "Knowledge, clusters, and competitive advantage". *Academy of management review* 29 (2): 258-271.

Tallman, S, Chacar, A. 2011 Knowledge Accumulation and Dissemination in MNEs: A Practice-Based Framework, *Journal of Management Studies* 48:2 March 2011

Teece, D., Pisano, G., Shuen, A., 1997, "Dynamic capabilities and strategic management". *Strategic Management Journal* 18 (7): 509-533.

Uzzi, Brian, 1997, "Social structure and competition in interfirm networks: The paradox of embeddedness". *Administrative Science Quarterly* 42 : 35-67.