Industrial and Corporate Change, Volume 16, Number 1, pp. 89–130
doi:10.1093/icc/dtl034
Advance Access published May 18, 2007

Desperately seeking spillovers? Increasing returns, industrial organization and the location of new entrants in geographic and technological space*

Barak S. Aharonson, Joel A. C. Baum and Maryann P. Feldman

Using detailed data on Canadian biotechnology firms during the 1990s, we explore the geographic scope of knowledge spillovers and the balance spillover-seeking and expropriation-avoidance in entrants' locations. Our findings indicate that knowledge spillovers are highly localized, with entrants attracted to incumbents' R&D employees and spending within 500 m, but not further. We also find that two local contextual factors enhance the tendency toward spillover seeking. One is increasing returns to positive information externalities that accompany concentrations of technologically similar firms. The other is the entrepreneurial and open industrial organization that arises when incumbents with direct ties to universities concentrate geographically. Our findings provide empirical evidence of forces promoting geographically concentrated and technologically specialized industrial micro-clusters, as well as factors reinforcing the significance of co-location for the creation of new knowledge.

1. Introduction

Although the benefits associated with industrial clustering, building on Marshall (1920), are well known, non-pecuniary externalities known as knowledge spillovers remain controversial, in part, because of the broad range of localized knowledge

*This research was supported in part by the Merck Frosst Canada & Co. Research Award on Canadian Competitiveness. We are grateful to Lee Fleming for helpful conversations, and for feedback from participants at the Danish Research Unit for Industrial Dynamics (DRUID) conference, the Center for Economic Policy Research (CEPR) conference on Integration and Technological Change, seminar participants at the Harvard Business School and Kellogg Graduate School of Management at Northwestern University, and an anonymous *ICC* reviewer for their helpful comments on earlier versions of this article. Whitney Berta, Jack Crane, Igor Kotlyar, Danny Tzabbar and Haley Waxberg all provided expert help with data collection and coding.

flows the term has come to encompass (Breschi and Lissoni, 2001). Yet, because technological knowledge has incomplete property rights and because geographic proximity provides an advantage for observation, interaction, collaboration, and inquiry, the idea of local non-market knowledge externalities has been embraced by theories of industrial clustering in the new economic geography.

In this view, information pertaining to new technologies is held to diffuse within small concentrated areas of individuals actively involved in similar technical problems through informal social and professional networks that emerge (Saxenian, 1994; Sorenson, 2003; Owen-Smith and Powell, 2004). In contrast to labor market pooling and specialized intermediate inputs, which are club goods that operate at a cluster or regional level, knowledge spillovers depend on face-to-face interactions between active researchers and are expected to be highly localized and to decay rapidly with distance (Rosenthal and Strange, 2003).

If knowledge spillovers matter for innovation and success in an industry, then they should be particularly influential in the location decisions of new entrants. Because knowledge spillovers lower the cost of innovating, resource-constrained startups should be attracted to incumbents that are the sources of economically useful knowledge. Studies showing higher entry rates in regions where activity is most concentrated are often taken as evidence of knowledge spillovers, but also consistent with benefits of labor market pooling and input sharing (e.g., Beaudry, 2001; Dumais *et al.*, 2002). The distance-sensitivity of knowledge spillovers suggests that they should be distinguished from agglomeration benefits such as labor market pooling and input sharing by the production of more localized effects. Corroborating this idea, recent studies show that entrants are sensitive to industry employment within a one mile radius and then attenuate rapidly (Rosenthal and Strange, 2003; Arzaghi and Henderson, 2004).

Although more persuasive evidence and a more nuanced understanding of the role of knowledge spillovers in the dynamics of industrial clustering are emerging from these recent micro-geographic analyses, they have not yet examined how entrants resolve the conundrum that by locating where they may benefit from other's ideas they also increase the risk of having their own ideas expropriated (Shaver and Flyer, 2000; Flyer and Shaver, 2003). Entrants that depend on innovation for their success and survival thus not only face strategic decisions about organizing their own R&D activities, but also how co-locating with similar firms affects their productivity.

We examine two dynamics that may moderate the tension between appropriation and expropriation, and thus strengthen or weaken the impact of localized knowledge spillovers on the intensity of spillover-seeking. One is increasing returns to positive information externalities that arise when firms in a given technological specialization concentrate in a small geographic area (e.g., Lucas, 1988; Romer, 1986, 1990). The other is industrial organization, with more entrepreneurial and open micro-environments fostering the development of social networks and norms of cooperation that facilitate the sharing, exchange and integration of information and knowledge (e.g., Saxenian, 1994; Brown and Duguid, 2000; Kogut, 2000).

We examine the influence of localized knowledge spillovers, increasing returns, and industrial organization on entrants' locations in the Canadian biotechnology industry. We exploit a detailed longitudinal data set that includes comprehensive information on all firms operating in the Canadian biotechnology industry during the 1990s. The data includes information on each firm's postal code, which allows us to distinguish locations at increments of 250 m or less, permitting us to examine the rate of spatial decay of knowledge spillovers over short distances. In addition, while the term biotechnology is often used to describe an industry, it is more accurately a technology for manipulating microorganisms that, over time, has been applied to a range of industrial sectors including agriculture, aquaculture, food and beverage, and pharmaceuticals. As a result of the cumulativeness of technological advances and the specificity of scientific novelty, potential knowledge spillovers should be stronger within than across to these distinct applications. Our data thus also permit us to examine new entrants' locations vis-à-vis incumbents within relatively narrow, specialized technological applications.

2. Co-location and seeking spillovers

Entrants need to be strategically minded as choices made at the time of startup influence future performance (Delmar and Shane, 2003). Entrepreneurship scholars suggest, however, that entrants do not typically choose where to locate their new businesses strategically but instead tend to establish them where they happen to be located at the time of founding (e.g., Shane, 2002). Evidence appears to support this claim at the regional level; for example, a person living or working in a given city is more likely to found a new enterprise in that city than in other locations (e.g., Buenstorf and Klepper, 2005). However, *within* a given area, entrants still face the decision as to precisely where to establish operations (Baum and Haveman, 1997; Stuart and Sorenson, 2003; Kalnins and Chung, 2004).

Such location choices are acutely relevant to entrants in knowledge-intensive industries that, lacking the resources and experience of more established firms, must rely more on external resources and knowledge to augment their own innovative capabilities (Aldrich and Auster, 1986; Baum and Oliver, 1991; Feldman, 1994), and lower their cost of innovation (Feldman, 2003). Opportunities to benefit from such knowledge spillovers depend on the spatial distribution of R&D activity, with more concentrated areas of inventive activity representing more promising locations. Diffusion of knowledge tends to be particularly localized for rapidly evolving technologies characterized by relatively high degrees of novelty and complexity that cannot be easily communicated, or codified into blueprints, contracts and journal articles (Audretsch and Feldman, 1996; Mowery & Ziedonis, 2001;

Almeida *et al.*, 2003). When knowledge is at its earliest stages of development, the ability to recognize potential applications of and effectively transfer new knowledge requires familiarity and understanding that develop through relationships and frequent face-to-face contact (Saxenian, 1994; Jaffe, 1989; Almeida and Rosenkopf, 2003; Owen-Smith and Powell, 2004).

When people with common technical interests concentrate geographically, dense local social and professional networks emerge as their close proximity leads them to encounter one another more frequently, both by chance and through local institutions, and to develop ties that are more likely to endure than more costly-to-maintain distant ties. By facilitating repeated interactions and development of overlapping social and professional connections, local concentrations of people engaged in similar technical activities create an environment facilitating trust-building and rapid and effective diffusion of ideas (Liebeskind *et al.*, 1996; Brown and Duguid, 2000; Kogut, 2000). Through these networks flows information about promising new technical developments and important unsolved puzzles that can stimulate innovation by facilitating novel combinations of ideas and technologies and identifying emerging market opportunities (Sorenson, 2003; Stuart and Sorenson, 2003; Owen-Smith and Powell, 2004).

If entrants seek to benefit from potential knowledge spillovers within such networks, they should tend to locate close to incumbents that are active in the process of invention—those spending more on R&D and employing more R&D workers—and so representing greater potential sources of new, high quality knowledge. The greater the level of nearby incumbents' inventive activity, the greater the pool of new knowledge into which an entrant can tap. The attractive force of incumbents' inventive activity will depreciate rapidly over distance, however, as informal social and professional networks become sparser and possibilities for face-to-face interactions between active researchers rarer (Owen-Smith and Powell, 2004).

It is not only close proximity to more inventive firms that enhances potential spillover benefits, however. Technological proximity also matters. The cumulativeness of technological advances and specificity of knowledge bases to particular technical areas and market applications makes the value of potential spillovers greater within rather than across specialized technological applications. Entrants' are better able to appreciate and absorb relevant new external information within their own technological specialization and market focus (Cohen and Levinthal, 1990). Moreover, within such technical specializations local social and professional networks are likely to be denser and more effective in providing access to relevant information.

Thus, while agglomeration benefits related to labor pooling and input sharing are likely to depend on the overall size of a region, knowledge spillover benefits depend on close proximity to specialized firms that share a relatively narrow common scientific or technological knowledge base (Surico, 2003). Compact neighborhoods that contain concentrations of knowledge-intensive resources will be the locus of knowledge spillovers. This is particularly true for rapidly evolving technologies for which meaning, context, and vocabulary are being defined. The foregoing ideas suggest the following hypothesis:

> *H1: The attraction of entrants to the locations of inventively active incumbents in their technological specialization declines rapidly with distance.*

3. Co-location and avoiding expropriation

Knowledge spillovers are not unidirectional, however (Shaver and Flyer, 2000; Flyer and Shaver, 2003). Knowledge has characteristics of a public good in that it is non-rival and non-excludable: once knowledge is created it is often difficult to contain and to prevent others from benefiting from or expropriating its value. Thus, once created, knowledge may spill over to benefit others able to monitor, observe, and recognize its potential. Entrants may suffer expropriation of their ideas and know-how if their knowledge spills over to neighboring incumbents' benefit. Yoffie (1993), for example, noted that semiconductor firms, concerned that their technology might spillover to competitors, were often wary of locating close to competitors.

Consideration of such strategic interactions complicates location decisions, suggesting that spillover-seeking may be tempered in technology and knowledge-intensive industries by a fear of expropriation. Although the literature suggests that localized competition may increase inventive output (Porter, 1990; Feldman and Audretsch, 1998), the idea that entrants would deliberately co-locate with competitors is only plausible if the net effects are positive. A key strategic issue for entrants is thus to balance potential costs of expropriation against likely gains from knowledge spillovers. Entrants who depend on rapid innovation for success and survival face strategic decisions not only about organizing their own R&D activities but also how locating relative to competitors affects their productivity and appropriability. Two factors seem acutely germane to specifying conditions affecting the balance of likely gains from knowledge spillovers relative to the risk of expropriation, and thus the intensity of local clustering.

3.1 Local increasing returns

The cumulative nature of invention manifests itself not just at firm and industry levels, but also at the geographic level, creating advantages for firms located in areas of concentrated inventive activity, and leading invention to exhibit pronounced geographical clustering (Lucas, 1988; Romer, 1986, 1990). These factors generate positive feedback loops or virtuous cycles as the increasing concentration of firms creates beneficial scale and information externalities by attracting additional labor and other inputs and further facilitating the exchange of ideas (Porter, 1990; Krugman, 1991). As spatial concentration of firms within a technological specialization increases, more ideas are generated for local diffusion and the realization of ideas is facilitated (Baptista and Swann, 1998, 1999; Wallsten, 2001), lowering the cost of innovation (Feldman, 2003). These economies increase with the concentration of technologically similar firms, attracting additional entrants, and enhancing the appeal of a location to still more entrants.

Consistent with these ideas, in addition to early studies showing that productivity increases substantially with increasing regional industry scale (e.g., Henderson, 1986), several more recent studies have shown that firms located in regions that were disproportionately concentrated in their own broadly defined (2-digit) industries tended to grow faster and produce more innovation when compared to more isolated firms located in regions that were concentrated in (2-digit) industries other than their own (Baptista and Swann, 1998; Beaudry and Breschi, 2003).

By enhancing knowledge externalities, increasing geographic concentration of technologically similar firms is expected to mitigate expropriation concerns. Within areas disproportionately concentrated by firms in a particular technological specialization, we therefore expect the greater magnitude of potential spillovers will weaken expropriation fears, leading to greater co-location of entrants with incumbents:

> *H2: The attraction of entrants to the locations of inventively active incumbents in their technological specialization increases with the incumbents' concentration.*

3.2 Local industrial organization

The productivity of a local environment depends not only on the scale of available inputs, but also on the way such inputs are organized. Saxenian (1994) argues that despite similarity in their local knowledge, labor, and input market characteristics, Silicon Valley outperformed Route 128 as a result of differences in their industrial organization and culture. Her ideas extend early work by Chinitz (1961) and Jacobs (1969) who argued that urban efficiencies depended importantly on the nature of urban interactions. Saxenian portrays Route 128's industrial organization and culture as being rigid and hierarchical and Silicon Valley's flexible and entrepreneurial.

In a recent study examining the different agglomerative effects of small and large firms, Rosenthal and Strange (2003) found empirical support for such an industrial organization effect. Their idea was that small firms are more likely to be entrepreneurial and open to interacting with their neighbors, with a greater agglomeration effect of small incumbents being predicted as a result. Consistent with their premise, locations with larger concentrations of small incumbents were found to be more attractive to new entrants. More generally, younger and smaller firms may tend to be more entrepreneurial (Hannan and Freeman, 1977, 1984; Chen and Hambrick, 1995), and lacking the resources, experience, and local network connections of more established incumbents, more open to interaction to access external resources and knowledge (Stinchcombe, 1965; Aldrich and Auster, 1986;

Baum and Oliver, 1991). Their lack of resources and R&D experience may also limit the threat of expropriation they pose to entrants. Younger and smaller incumbents may also tend to be more willing to interact with entrants because, sharing similar attributes, their deeper insight into each other's situations and behaviors enables them to engage in activities demanding high-trust kinds of cooperation (Zucker, 1986; Ring and Rands, 1989).

Alliances between incumbents and universities signal an additional element of local industrial organization (Porter and Stern, 2001). These alliances serve to diffuse university norms of open science and collaboration, fostering a local culture of learning and norms of cooperation that reinforce the operation of social and professional networks that serve as channels for the exchange, diffusion, and integration of information and ideas (Owen-Smith and Powell, 2004). These alliances also play a critical role in bridging technology and industry, serving as channels through which specialized knowledge is created and scientists and industry can work together on product research, development, and commercialization. Indeed, Zucker *et al.* (1998) found that spillovers appear to explain the knowledge transfer process only when formal contractual relationships between university scientists and firms are ignored. Alliances between incumbents and universities may thus greatly reduce expropriation concerns.

These ideas and findings regarding the influence of local industrial organization suggest that by enhancing knowledge externalities, inventively active incumbents that are younger and smaller and/or allied with universities trigger fewer expropriation concerns, leading to greater co-location of entrants with the incumbents. Therefore:

> *H3: The attraction of entrants to the locations of inventively active incumbents in their technological specialization increases with the incumbents' (1) youth and smallness and (2) university alliances.*

4. Data and methods

We tested our hypotheses using data on the 675 biotechnology firms operating in Canada at any time between January 1991 and December 2000 [255 (38%) exited the industry before the end of 2000]. The sample included 206 entrants established during the period of January 1991 through December 2000. We compiled our data using *Canadian Biotechnology*, an annual directory of Canadian firms active in the biotechnology field published since 1991. *Canadian Biotechnology* is the most comprehensive historical listing of Canadian biotechnology firms, providing information on their technological specializations, R&D employment and spending, products, alliances, and locations.[1] We cross-checked this information with

[1]The technological specializations are: agriculture, aquaculture, horticulture, forestry, engineering, environmental, food, beverage and fermentation, veterinary, energy, human diagnostics, human therapeutics, human vaccines, biomaterials, cosmetics, and mining.

The Canadian Biotechnology Handbook (1993, 1995, 1996), which lists information for a more restrictive set of *core* firms that are entirely dedicated to biotechnology.

We measured each firm's location by converting its six-character postal code address into latitude and longitude measurements. The form of the postal code is ANA NAN, where A is an alphabetic character and N is a numeric character. The first character of a postal code represents a province or territory, or a major sector entirely within a province. The first three characters of the postal code identify the Forward Sortation Area (FSA). Each FSA is associated with a single postal facility from which mail delivery originates. The average number of addresses within an FSA is approximately 7000. The last three characters of the postal code identify the Local Delivery Unit (LDU). Each LDU represents one or more mail delivery points, with the average number of addresses within an LDU about 15.[2] As of May 2001, there were over 1600 FSAs encompassing more than 750,000 LDUs.

Figure 1 shows the geographic distribution of biotechnology firms in Canada for the years 1991 and 2000. Overall, the industry is highly clustered within a small number of compact regions. The regions, labeled by the province (city) in which the majority of firms are located, are: British Columbia (Vancouver), Alberta (Calgary), Alberta (Edmonton), Saskatchewan (Saskatoon), Manitoba (Winnipeg), Ontario (Toronto), Ontario (Ottawa), Quebec (Montreal), Quebec (Quebec City), New Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland.

Use of postal codes has several advantages. First, it permits us to examine agglomeration effects within compact geographic areas comparable to or smaller in size than recent studies (e.g., Wallsten, 2001; Rosenthal and Strange, 2003; Stuart and Sorenson, 2003; Arzaghi and Henderson, 2004; Duranton and Overman, 2005), and much smaller than US zip codes, political or administrative jurisdictions such as states and counties or statistical units such as MSAs (Metropolitan Statistical Area) and SMSAs (Standard Metropolitan Statistical Area) used in most agglomeration research.[3] Second, it allows us to avoid well-known problems of aggregation (see Cressie, 1991) by permitting use of distances between firms rather than counts within arbitrary jurisdictions. Third, it permits us to combine elements of standard point pattern analysis from quantitative geography (e.g., Gatrell *et al.*, 1996; Cressie, 1991; Durnaton and Overman, 2005), which do not lend themselves to dynamic, causal analysis, with regression techniques that do.

[2]In urban areas, a postal code corresponds to one of the following: one block-face (i.e., one side of a city street between consecutive intersections); a Community Mail Box; an apartment or office building; a large organization or federal government department, agency or branch.

[3]Stuart and Sorenson (2003) report that the average zip code in their study of biotechnology firm founding is 27.4 square miles (44.4 km^2). MSAs are larger still, with the mean area of an MSA in the United States equal to 10,515 square miles (17,042 km).

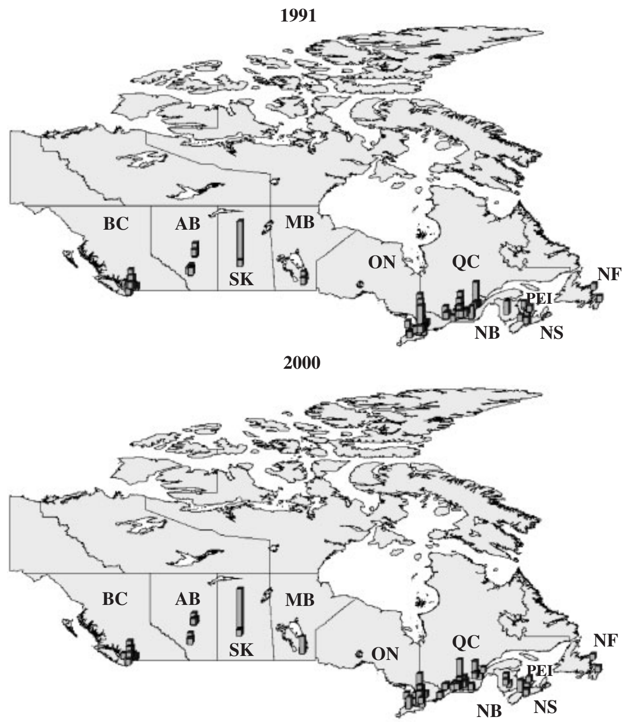

Figure 1 Geographic clustering of biotechnology firms in Canada.

4.1 Dependent variable and model specification

The dependent variable for our analysis is a dichotomous variable indicating whether or not a biotechnology firm entered a postal code in a given year. Of course, there are many postal codes (over 750,000), only a small fraction of which actually experience an entry, and many of which are not plausibly at risk of experiencing an entry. Accordingly, we adopted a case-control approach and employed rare events logistic regression to obtain our estimates (Lancaster and Imbens, 1996; King and Zeng, 2001; Tomz, 2003).

Since the postal codes in which entries occurred provide the most information regarding factors influencing entrants' locations, in each observation year, all postal codes in which an entry occurred were included in the analysis (Lancaster and Imbens, 1996; King and Zeng, 2001). We combined this sample of "entry postal codes" with a random sample of matched "non-entry postal codes" in which no entry occurred. These non-entry postal codes—the control group—must be "at risk" of entry and similar in character to entry postal codes, for example, matched in terms of geographic location and demographic features (e.g., population density and land use).

We matched each entry postal code with a random sample of postal codes from the same FSA as the entry postal code, but did not experience an entry that year.

Thus, for an entry in postal code "M4K 1S2" in 1995, we drew a random sample of postal codes beginning with "M4K" that did not experience an entry in 1995. This sampling frame for the non-entry postal codes ensures their similarity to the entry postal codes in terms of location and other important features, as well as the plausibility of their being at risk of entry. Because the information contributed by adding case controls is marginally decreasing, using this procedure, we matched each entry with five non-entry postal codes (King and Zeng, 2001; Jensen, 2003). Thus, a control group of 1030 randomly sampled non-entry postal codes was added to the case sample of 206 entry postal codes, which resulted in a final sample of 1236 observations.

Although the case control approach helps solve sampling problems associated with the analysis of rare events, it creates an estimation problem. Specifically, logistic regression tends to underestimate effects of factors predicting positive outcomes (here entries) for samples in which the proportion of positive outcomes is different from the proportion in the population. King and Zeng (2001) showed, however, that weights can be used to correct the estimation bias created by oversampling rare events. We employed Tomz's (2003) ReLogit (version 2.0) Stata procedure, which implements King and Zeng's weighting approach, to estimate a rare events logistic regression model. In implementing the model, we assumed that the proportion of positive outcomes (entries) in the population is 0.003. We derived this estimate by dividing the number of entries (206) by the total number of postal codes in the FSAs containing the entry postal codes (70,166). The proportion is 0.167 (i.e., 1 : 5) in the sample for the analysis. We estimated our models with robust standard errors based on the Huber/White/sandwich estimator of variance and grouped according to the FSA given the likely non-independence of entries within FSA.

4.2 Independent variables

We measured the intensity of incumbents' inventive activity using two indicators: R&D spending (in 1991 Canadian dollars) and number of R&D employees. To test H1, we summed (separately) the value of R&D spending and number of R&D workers of all incumbents in the entrant's technological specialization within rings of 0 to 500 m, 500 m to 2 km, and 2 to 10 km from the entrant's postal code.[4] For each of the non-entry postal codes matched to an entry, we computed the same values (for the matched entrant's technological specialization). These variables were logged to normalize their distributions, and to avoid simultaneity problems were computed based on incumbents' inventive activity in the year prior to the entry. Positive coefficients for these variables indicate that entrants tend to co-locate with inventively active incumbents in their technologically specialization supporting H1.

[4]In preliminary analyses, we examined rings of 500 m and 1 km increments for all our variables and found that this set of rings provided the most parsimonious and efficient estimates.

Using this ring approach, which is now standard in location studies (e.g., Wallsten, 2001; Rosenthal and Strange, 2003; Arzaghi and Henderson, 2004), we can estimate empirically the extent of geographic decay in agglomeration effects.[5]

To test H2, we computed the proportion of all Canadian incumbents in the entrant's technological specialization located within rings of 0 to 500 m, 500 m to 2 km, and 2 to 10 km from an entrant's postal code as well as the non-entry postal codes matched to the entry. We computed these variables based on incumbents' locations in the year prior to the entry to avoid simultaneity problems. We then interacted these variables with the inventive activity variables for their respective rings. Positive coefficients for the interactions, indicating that the agglomerative effect of incumbents' inventive activity is stronger in locations where there is a high concentration of technologically similar firms, would support H2.

To test H3, we computed three sets of inventive activity variables to capture the demography of local industrial organization. The first set disaggregated the composite inventive activity variables used to test H1 according to incumbents' age, the second according to incumbents' size, and the third for incumbents with and without a university alliance. For age, inventive activity was computed separately for incumbents aged 1–2 years, 3–5 years, and over 5 years. For size, they were recomputed for incumbents with 1–10 R&D employees, 11–50 R&D employees, and over 50 R&D employees.[6] Support for H3 requires larger positive (or smaller negative) coefficient estimates for the inventive activity of younger and smaller incumbents, and for incumbents with university alliances.

[5]Although the sampling frame we adopted ensures the similarity of entry and non-entry locations, it also results in their proximity, with the average distance between entry and matched non-entry locations in our sample being 4.4 km. This poses a challenge to estimating location effects because the ring variables for entry and non-entry locations may overlap, with the potential for overlap increasing with the distance of the ring. Given the average 4.4 km distance for our sample, the 0 to 500 m and 500 m to 2 km rings pose limited concern in this regard (on average, they are not overlapped), but the 2–10 km rings overlap roughly 50%, on average. This shared variance may result in a downward bias in coefficients estimated for the 2–10 km rings, and thus an upward bias in the estimated rate of geographic decay. We addressed this potential bias in three ways. First, we estimated more fine-grained 500 m and 1 km ring increments and these revealed the same decay rate as the aggregate measures we report. Second, all models include 'entry group' fixed effects; that is, a fixed effect for each entry and its five matched non-entry locations. Third, we compare estimated rates of geographic decay for technologically similar incumbents' inventive activity with those for proximity to universities, which we expect (and find) to be much slower. We are grateful to an anonymous reviewer for altering us to this measurement issue and his/her suggestions for addressing it.

[6]In preliminary analyses, we examined more fine-grained age and size categories and found that these composite categories provided the most parsimonious and efficient estimates.

4.3 Control variables

A range of additional factors is likely to influence entrants' locations; accordingly, we control for location, technology, and year fixed effects as well as variables for key local characteristics.

We included fixed effects for both "entry group" (i.e., each entry and its five non-entry control locations) and region (as defined in Figure 1) in our model specification. These fixed effects control for a wide range of local and regional features that might influence entry locations. At the entry group level, such features include local variation in human population density, land use, rent, and property values. Entry group fixed effects also control for the possibility that a particular location is more attractive because resources have been invested in an effort to attract high technology firms to it, for example, by establishment of an incubator designed to elicit spillovers. At the regional level, they include variation in urbanization, labor market pooling and costs, access to shared and complementary inputs including venture capital, and potential entrants.

We also included fixed effects for technological specialization to control for differences in the location patterns of entrants in across specializations, which vary in their technological maturity and required labor skills and complementary resources. Finally, we included year fixed effects to control for potential changes in the propensity of entrants to co-locate over time.

We also controlled for two sets of variables capturing more fine grained environmental features likely to influence entrants' locations. The first set controls for inventive activity of incumbents in *other* technological specializations than the entrant. While entrants' greater ability to appreciate and absorb external information relevant to their own specialization heightens spillover benefits and expropriation risks within each specialization, the inventive activity of incumbents in other specializations may still impact entrants' location choices as a result of similarities in labor, shared and complementary resources, and other inputs across biotechnology subfields. To control for this possibility, we computed variables analogous to those used to test H1, but based on incumbents' inventive activity in technological specializations other than the entrant's. We expect that the effects of incumbents' inventive activity on entrants' locations will be stronger and more geographically localized effects within technological specializations than across them.

The second set controls for the location of universities. Zucker *et al.*'s (1998) work on the clustering of US biotechnology firms around universities (and 'star scientists' in particular), as well as earlier studies demonstrating that proximity to universities enhances firms' inventive output (e.g., Jaffe, 1989; Feldman, 1994), suggest that entrants may co-locate with universities (see also Agrawal and Cockburn, 2003). This may occur both because entrants are attracted to the scientists and other resources affiliated with universities (e.g., there may be dedicated buildings receiving public subsidies to host new entrants), and because university scientists in the biotech-related sciences may tend to establish their own biotechnology firms in the immediate vicinity of their universities (Stuart and Sorenson, 2003). Scientists may, for example, want to be able to commute easily between their university and business offices. Former students may decide to locate start-ups close to their former university because they like it there. We therefore control for the location of universities to ensure that the estimates for incumbents' inventive activity do not spuriously capture a tendency for entrants to co-locate with universities. Because allied universities and firms tend to be co-located (e.g., Zucker *et al.*, 1998), this control is also necessary to distinguish the effects of incumbents' university alliances on agglomeration from proximity to universities. The control variables counted the number of universities within rings of 0 to 500 m, 500 m to 2 km, and 2 to 10 km from the entrant's postal code and random non-entry postal codes matched to the entry.

Tables A1 and A2 present descriptive statistics and bivariate correlations for the independent and control variables included in the analysis. Table A1 gives the information for inventive activity variables computed based on incumbents' R&D employees; A2 for those computed based on their R&D spending. The tables give statistics for disaggregated versions of the entrant's specialization inventive activity variables for the <500 m ring. The concentration and university ring variables are identical in both tables; but their correlations with the R&D spending and employee inventive activity variables differ.

Summary statistics are given separately for entry postal codes and non-entry postal codes. Notably, while for rings beyond 500 m there are few significant differences in variable means for entry and non-entry postal codes, for the <500 m ring *all* means are significantly larger for the entry than non-entry postal codes ($P < 0.05$ or greater). Thus, the random sample of non-entry postal codes is indistinguishable from the entry postal codes beyond the 500 m ring, but despite being drawn from the same FSA, significantly different within it. This pattern of differences in means suggests a good match between case and control samples, as well as a strong localization of agglomeration effects.

While, as expected, correlations between composite and disaggregated versions of the inventive activity variables are moderately high, because the composite and disaggregated variables are not included in the same models, this poses no estimation problem. Correlations among the age, size, and alliance category variables, which are included in the same models, are substantially lower. More generally, correlations among the variables are low to moderate and so do not pose an estimation problem.

4.4 Initial results

Table 1 reports initial empirical estimates testing the basic spillover-seeking hypothesis (H1), controlling for inventive activity of incumbents in *other* technological specializations than the entrant, university locations, and entry group, region, technological specialization, and year fixed effects. Model 1a shows

Table 1 Rare event logistic regression models of biotech firm entry—spillover seeking and avoidance

	Inventive activity: R & D employees						Inventive activity: R & D expenses					
	Model 1a			Model 1a(2)			Model 1b			Model 1b(2)		
	β	S.E.	*P*	*β*	S.E.	*P*	*β*	S.E	*P*	*β*	S.E.	*P*
Inventive activity—entrant's specialization												
<500 m	0.785	0.188	***	0.771	0.244	**	0.162	0.037	***	0.208	0.055	***
<500 m^2				−0.005	0.026					−0.002	0.001	+
500 m to 2 km	0.279	0.153	*	0.278	0.160	*	0.033	0.026		0.033	0.026	
2 to 10 km	−0.051	0.043		−0.052	0.045		−0.014	0.009	+	−0.014	0.009	+
Inventive activity—other specializations												
<500 m	0.358	0.062	***	0.656	0.120	***	0.086	0.013	***	0.128	0.024	***
<500 m^2				−0.023	0.008	**				−0.001	0.000	**
500 m to 2 km	−0.018	0.038		−0.021	0.039		0.001	0.008		0.000	0.008	
2 to 10 km	−0.024	0.019		−0.019	0.019		−0.002	0.004		−0.002	0.004	
Number of universities												
<500 m	3.858	0.805	***	3.729	0.801	***	3.641	0.751	***	3.583	0.757	***
500 m to 2 km	0.925	0.566	+	0.997	0.527	*	1.110	0.527	*	1.167	0.535	*
2 km to 10 km	0.355	0.298		0.420	0.301	+	0.560	0.349	+	0.651	0.350	*
Fixed effects												
Entry group	incl.			incl.			incl.			incl.		
Region	incl.			incl.			incl.			incl.		
Technological specialization	incl.			incl.			incl.			incl.		
Year	incl.			incl.			incl.			incl.		
Log likelihood	−448.30			−442.99			−447.38			−443.90		
Likelihood ratio test				10.620	* (2 df)					6.960	* (2 df)	

Notes: *P*-level: + <0.01; *<0.05; **<0.01; ***<0.001; the sample included 1236 0bservations (206 entry postal codes; 1030 random non-entry postal codes); Incl. Included.

the effects of incumbents' inventive activity in the entrant's specialization indexed by their R&D employment; Model 2a shows the same model but with incumbents' inventive activity indexed by their R&D spending.

The effect of incumbents' inventive activity is strikingly local and strongly agglomerative. The effect of technologically similar incumbents' inventive activity on entrants' location, measured as R&D employees, is significant for both the <500 m and 500 m to 2 km rings, but the effect is nearly three times larger within the smaller ring. A comparison-of-means test (Wonnacott and Wonnacott, 1970) indicates that this difference is statistically significant ($P < 0.001$). Measured as R&D spending, the effect of incumbents' inventive activity is limited to the <500 m ring. Also notable is the negative coefficient for technologically similar incumbents' R&D spending for the 2–10 km ring, which is marginally significant ($P < 0.10$). This indicates that, in addition to being more likely to enter locations within 500 m of R&D activity in their specialization, entrants are less likely to enter locations more than 2 km from such activity.

The effects of technologically dissimilar incumbents' inventive activity are similarly localized. However, despite the similarly strong localization of the agglomerative effects of incumbents' inventive activity in the entrant's own and other technological specializations, the effect for the entrant's specialization is roughly twice as strong for both R&D employees and R&D spending. Comparison-of-means tests indicate that these differences are statistically significant ($P < 0.01$). The effect of universities is also highly localized, with the effect three to four times stronger within 500 m than between 500 m and 2 km, but the decay is slower, with evidence that university agglomeration effects extends as far as 10 km.

These initial estimates strongly support H1. To examine the robustness of this result, we re-estimated Models 1a and 1b using quadratic specifications for the <500 m ring inventive activity variables to allow for possible congestion or crowding effects.[7] Likelihood ratio tests in Table 1 indicate that the quadratic specifications improve on the linear ones, with the squared terms significant and negative for both other specialization variables, and weakly so for the same specialization variable in the R&D spending model ($P < 0.10$). There are thus limits to the local agglomerative effects of incumbents' inventive activity, particularly in other specializations.

Figures 2a and 2b plot multipliers estimated from Models 1a(2) and 1b(2) that compare the likelihoods of entry in a postal code with no inventive activity within 500 m and postal codes with inventive activity levels up to two standard deviations above the mean. A multiplier of two indicates a doubling of the likelihood, of three tripling, and so on. As the figures show, the agglomerative influence of incumbents' inventive activity is large in magnitude, particularly within the entrant's specialization, and in addition, that the effect of R&D employees is roughly double that of

[7]Quadratic specifications for the other rings were not significant.

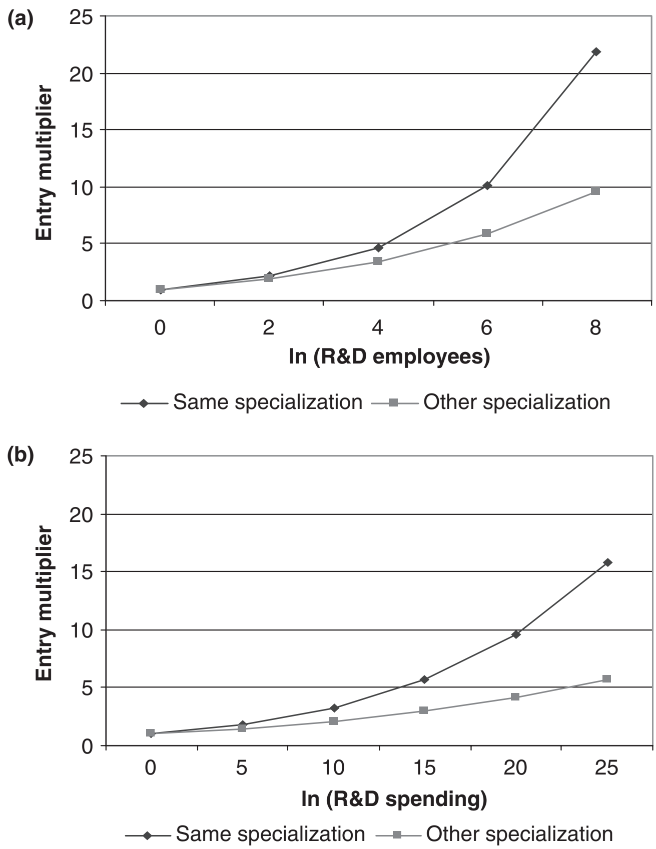

Figure 2 Multipliers for R&D employees within 500 m.

R&D spending. As the figures also show, none of the quadratic effects reaches its maximum within plotted ranges of inventive activity; the negative squared terms thus decrease the marginal effect of increasing inventive activity on agglomeration but do not reduce it.

4.5 Increasing returns

Table 2 extends the analysis by including the entrant's specialization concentration variables and their interactions with inventive activity. These models test the increasing returns hypothesis (H2), which predicts that spillover-seeking in a given technological specialization will be stronger in areas concentrated in that specialization. Given the findings in Table 1, we focus on the effects of concentration of the entrant's specialization within 500 m. Likelihood ratio tests reported in Table 2 indicate that including the concentration effects improves model fit significantly for both measures of inventive activity.

Table 2 Rare event logistic regression models of biotech firm entry—controlling for increasing returns

	Model 2a			Model 2b		
	R&D employees			R&D expenses		
	β	S.E.	*P*	β	S.E.	*P*
Inventive activity—Entrant's specialization						
<500 m	1.961	0.539	***	0.484	0.153	**
<500 m^2	−0.266	0.116	**	−0.016	0.008	*
500 m to 2 km	0.300	0.165	*	0.036	0.029	
2–10 km	−0.042	0.047	*	−0.012	0.010	
Concentration—entrant's specialization						
<500 m	74.841	25.664	**	47.465	19.277	**
× Inventive Activity <500m	−74.221	28.095	**	−13.493	3.864	***
× Inventive Activity <500m^2	11.260	4.884	**	0.539	0.241	*
Inventive activity—other specializations						
<500 m	0.653	0.124	***	0.128	0.025	***
<500 m^2	−0.022	0.008	**	−0.001	0.000	**
500 m to 2 km	−0.021	0.040		0.001	0.008	
2–10 km	−0.018	0.019		−0.002	0.004	
Number of universities						
<500 m	3.717	0.786	***	3.626	0.744	***
500 m to 2 km	0.923	0.570	+	1.096	0.537	*
2–10 km	0.396	0.293		0.616	0.352	*

Table 2 Continued

	Model 2a			Model 2b		
	R&D employees			R&D expenses		
	β	S. E.	P	β	S. E.	P
Fixed effects						
Entry group	incl.			incl.		
Region	incl.			incl.		
Technological specialization	incl.			incl.		
Year	incl.			incl.		
Log likelihood	−438.67			−436.95		
Likelihood ratio test versus models 1 a(2) and 1 b(2)	8.65	* (3 df)		12.07	* (3 df)	

Notes: *P*-levels: +<0.01; *<0.05; **<0.01; ***<0.001; the sample included 1236 observations (206 entry postal codes; 1030 random non-entry postal codes); Incl. Included.

When the main and interactive effects of concentration are added, the quadratic specifications for incumbents' inventive activity in the entrant's specialization become significant for both R&D employees (Model 2a) and R&D spending (Model 2b). The positive linear terms are consistent with spillover seeking, while negative squared terms suggest that crowding may dampen further clustering at higher levels of inventive activity. The concentration main effects are significant and positive, and concentration × inventive activity interaction terms significant in opposition to the main effects of inventive activity. Supporting H2, the interaction terms indicate that the agglomerative effects of incumbents' inventive activity are stronger in locations with a higher national share of incumbents in an entrant's technological specialization.

Figures 3a and 3b, which show multipliers estimated from Models 2a and 2b, illustrate these interaction effects. In the figure, the vertical (y) axis indicates the multiplier for the likelihood of entry relative to a postal code with no inventive activity and zero concentration, the front horizontal (x) axis indicates the concentration of the entrant's specialization within 500 m, and the right horizontal (z) axis indicates the level of inventive activity in the entrant's specialization within 500 m. The figures are plotted for levels up to two standard deviations above the means for inventive activity as well as concentration.[8]

As the figures show, when local concentration is low (1%), the likelihood of entry increases with initial increases in inventive activity, but then declines with further increases. The declines begin when the number of R&D workers within 500 m exceeds 55 (e^4) and R&D spending exceeds \$3.3 million (e^{15}). As the local concentration level increases beyond 3%, however, the impact of further increases in inventive activity shifts from a decline to a rise in the likelihood of entry. Thus, as predicted by H2, as local concentration increases, inventive activity induces greater clustering. These findings illustrate graphically how virtuous cycles of co-location can produce high levels of clustering in a small number of locations.

4.6 Industrial organization

Tables 3, 4 and 5 present models in which the agglomeration effects of incumbents' inventive activity vary by incumbent age, size, and university alliance status, respectively. The models in these tables test the industrial organization hypothesis (H3), which predicts that agglomeration will be greater in locations where younger, smaller, and university-allied incumbents in the entrant's specialization are engaged in inventive activity. Based on the estimates in Table 1, we again focus on incumbents' age, size, and alliance status within a 500 m radius, and also specify quadratic effects for each age, size, and alliance category. Because quadratic specifications for the >5 years age category in Table 3 and 1–10 R&D employees size category in Table 4 were

[8]The concentration scale in Figures 3a and 3b begins at 1% since incumbents' inventive activity levels for locations with concentrations below this level are severely restricted.

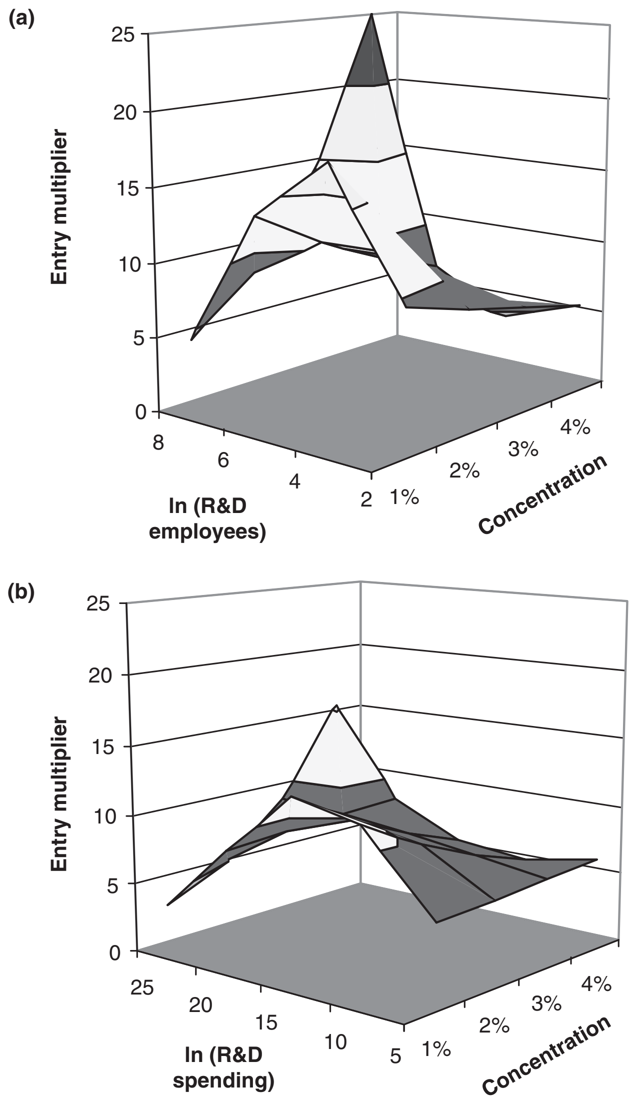

Figure 3 Multipliers for R&D employees within 500 m—concentration effects.

insignificant, we re-estimated these models specifying linear effects for these categories. In the reduced models, which are not significantly different from their respective full models, estimates for all other results are unchanged.

The estimates in Table 3 indicate that the inventive activity of adolescent and older, but not the youngest, incumbents have the strongest agglomerative effects. Inventive activity of 3–5 year old incumbents in the entrant's specialization affect agglomeration positively initially and negatively at higher levels of inventive activity.

Table 3 Rare event logistic regression models of biotech firm entry—controlling for incumbent age

	R&D employees						R&D expenses					
	Model 3a			Model 3a(2)			Model 3b			Model 3b(2)		
	β	S.E.	P	β	S.E.	P	β	S.E.	P	β	S.E.	P
Inventive activity—entrant's specialization												
500 m (Age 1–2 Years)	−2.393	1.204	*	−2.382	1.199	*	−1.399	0.700	*	−1.393	0.700	*
<500 m^2 (Age 1–2 Years)	0.760	0.325	**	0.757	0.323	**	0.103	0.051	*	0.103	0.051	*
<500 m (Age 3–5 Years)	3.565	0.816	***	3.560	0.813	***	0.621	0.157	***	0.620	0.158	***
<500 m^2 (Age 3–5 Years)	−0.490	0.155	**	−0.990	0.154	***	−0.019	0.007	**	−0.019	0.007	**
<500 m (Age > 5 Years)	0.961	0.544	*	0.866	0.191	***	0.122	0.102		0.153	0.037	***
<500 m^2 (Age > 5 Years)	−0.019	0.091					0.001	0.004				
500 m to 2 km	0.255	0.143	*	0.254	0.143	*	0.021	0.023		0.021	0.023	
2–10 km	−0.056	0.049		−0.057	0.049		−0.017	0.010	*	−0.017	0.010	*
Inventive activity—other specializations												
<500 m	0.764	0.124	***	0.763	0.124	***	0.135	0.023	***	0.135	0.023	***
<500 m^2	−0.029	0.008	***	−0.029	0.008	***	−0.001	0.000	**	−0.001	0.000	**
500 m to 2 km	0.005	0.039		0.005	0.039		0.006	0.003		0.006	0.008	
2–10 km	−0.020	0.019		−0.020	0.019		−0.002	0.004		−0.002	0.004	
Number of universities												
<500 m	3.915	0.768	***	3.919	0.768	***	3.837	0.745	***	3.830	0.744	***
500 m to 2 km	1.027	0.574	*	1.025	0.574	*	1.294	0.515	**	1.294	0.515	**
2–10 km	0.615	0.338	*	0.611	0.335	*	0.775	0.370	*	0.776	0.369	*

(continued)

Table 3 Continued

	R&D employees						R&D expenses					
	Model 3a			Model 3a(2)			Model 3b			Model 3b(2)		
	β	S. E.	*P*	β	S. E.	*P*	β	S. E.	*P*	β	S. E.	*P*
Fixed effects												
Entry Group	incl.			incl.			incl.			incl.		
Region	incl.			incl.			incl.			incl.		
Technological specialization	incl.			incl.			incl.			incl.		
Year	incl.			incl.			incl.			incl.		
Log likelihood	−428.19			−428.20			−433.58			−433.63		
Likelihood ratio test versus Models 3a and 3b				0.03 (1 df)						0.10	(2 df)	

Notes: *P*-levels: + <0.01; *<0.05; **<0.01; ***<0.001; the sample included 1236 observations (206 entry postal codes; 1030 random non-entry postal codes); Incl. Included.

Table 4 Rare event logistic regression models of biotech firm entry—controlling for incumbent size

	R&D employees						R&D expenses					
	Model 4a			Model 4a(2)			Model 4b			Model 4b(2)		
	β	S.E	*P*	*β*	S.E	*P*	*β*	S.E	*P*	*β*	S.E	*P*
Inventive activity—entrant's specialization												
<500 m (1–10 R&D employees)	−0.274	1.095		0.053	0.407		−0.055	0.135		−0.026	0.076	
<500rn Squared (1–10 R&D employees)	0.130	0.316					0.002	0.006				
500 m (11–50 R&D employees)	1.358	0.645	*	1.299	0.599	*	0.391	0.091	***	0.386	0.091	***
<500 m^2 (11–50 R&D employees)	−0.100	0.071	+	−0.087	0.057	+	−0.006	0.002	***	−0.006	0.002	***
<500 m (>50 R&D employees)	1.620	0.437	***	1.612	0.445	***	0.586	0.130	***	0.586	0.131	***
<500 m^2 (>50 R&D employees)	−0.136	0.072	*	−0.134	0.076	*	−0.013	0.005	**	−0.013	0.005	**
500 m to 2 km	0.260	0.149	*	0.269	0.154	*	0.030	0.023		0.031	0.023	
2–10 km	−0.042	0.046		−0.042	0.046		−0.010	0.011		−0.010	0.011	
Inventive activity—other specializations												
500 m	0.670	0.115	***	0.667	0.117	***	0.141	0.023	***	0.141	0.023	***
<500 m^2	−0.025	0.008	***	−0.025	0.008	***	−0.001	0.000	**	−0.001	0.000	**
500 m to 2 km	−0.024	0.040		−0.025	0.040		−0.002	0.007		−0.002	0.007	
2–10 km	−0.020	0.019		−0.020	0.019		−0.002	0.005		−0.002	0.005	
Number of universities												
500 m	3.925	0.785	***	3.905	0.783	***	4.062	0.762	***	4.057 0.760	0.760	***
500 m to 2 km	1.102	0.597	*	1.094	0.597	*	1.462	0.527	**	1.460 0.528	0.528	**
2–10 km	0.501	0.313	+	0.496	0.311	+	0.808	0.370	*	0.811	0.371	*

(continued)

Table 4 Continued

	R&D employees						R&D expenses					
	Model 4a			Model 4a(2)			Model 4b			Model 4b(2)		
	β	S.E	P	β	S.E	P	β	S.E	P	β	S.E	P
Fixed effects												
Entry group	incl.			incl.			incl.			incl.		
Region	incl.			incl.			incl.			incl.		
Technological specialization	incl.			incl.			incl.			incl.		
Year	incl.			incl.			incl.			incl.		
Log likelihood	−438.37			−438.46			−429.49			−429.51		
Likelihood ratio test versus Models 4a and 4b				0.18		(1 df)				0.05	(1 df)	

Notes: P-levels: + <0.01; *<0.05; **<0.01; ***<0.001; the sample included 1236 observations (206 entry postal codes; 1030 random non-entry postal codes); Incl. Included.

Table 5 Rare event logistic regression models of biotech firm entry—controlling for university alliances

	Model 5a			Model 5b		
	R&D employees			R&D expenses		
	β	S.E	*P*	β	S.E	*P*
Inventive activity—entrant's specialization						
<500 m (University alliance)	0.926	0.550	*	0.179	0.104	*
<500 m^2 (University alliance)	0.019	0.111		−0.001	0.003	
<500 m (No university alliance)	0.873	0.402	*	0.309	0.087	***
<500 m^2 (No university alliance)	−0.076	0.068		−0.012	0.003	***
500 m to 2 km	0.312	0.175	*	0.029	0.029	
2–10 km	−0.051	0.045		−0.015	0.010	+
Inventive activity—other specializations						
<500 m	0.651	0.121	***	0.125	0.024	***
<500 m^2	−0.023	0.008	**	−0.001	0.000	**
500 m to 2 km	−0.026	0.039		0.001	0.008	
2–10 km	−0.020	0.019		−0.002	0.004	
Number of universities						
<500 m	3.839	0.787	***	3.554	0.736	***
500 m to 2 km	1.107	0.513	*	1.116	0.532	*
2–10 km	0.537	0.312	*	0.602	0.345	*

(continued)

Table 5 Continued

	Model 5a			Model 5b		
	R&D employees			R&D expenses		
	β	S. E.	P	β	S. E.	P
Fixed effects						
Entry group	incl.			incl.		
Region	incl.			incl.		
Technological specialization	incl.			incl.		
Year	incl.			incl.		
Log likelihood	−441.30			−442.97		

Notes: *P*-levels: + <0.01; *<0.05; **<0.01; ***<0.001; the sample included 1236 observations (206 entry postal codes; 1030 random non-entry postal codes); Incl. Included.

For incumbents over 5 years of age, the effect is entirely positive. The quadratic effect for inventive activity of incumbents aged <2 years is significant for both R&D employees [Model 3a(2)] and spending [Model 3b(2)]. Notably, however, the linear terms for this age group are negative and squared terms positive. This means that entry is less likely in locations where the youngest incumbents in the entrant's specialization employ few R&D workers/spend little on R&D, and more likely where they employ many R&D workers/spend a lot on R&D. Although consistent with the prediction in H3 that young incumbents' inventive activity would attract entry, this support is overshadowed by the larger agglomerative effects of adolescent and older incumbents' inventive activity. Additionally, the negative linear effect dominates the positive squared effect over the observed range of inventive activity for young firms.

Figures 4a and 4b show this graphically, giving multipliers estimated from Models 3a(2) and 3b(2) that compare the likelihoods of entry in a postal code with no inventive activity by incumbents in a given age category and postal codes with inventive activity levels in that age category up to two standard deviations above the mean. The figures show large agglomerative influence of adolescent incumbents' (aged 3–5 years) inventive activity. The figures also show that the effect of adolescent incumbents' R&D employees begins to decline when their employment exceeds 20 (e^3) R&D workers within a 500 m radius, and that the marginal effect of R&D spending over \$1.2 million (e^{14}) is small. This suggests that entrants may view inventively active adolescent incumbents in their specialization as potential expropriation threats. Overall, however, and in contrast to H3, young inventively active incumbents are a weak attractive force.

The estimates in Table 4, which disaggregate incumbents' inventive activity by size category, again fail to support H3. The inventive activities of large and medium-sized incumbents in an entrant's specialization have large agglomerative effects; small incumbents' inventive activity has no effect. Figures 5a and 5b chart these effects graphically, giving multipliers estimated from Models 4a(2) and 4b(2) that compare the likelihoods of entry in a postal code with no inventive activity in a given size category and postal codes with inventive activity levels in that size category up to two standard deviations above the mean. The figures show the greater agglomerative influence of large incumbents' inventive activity, as well as the decrease in attractiveness implied by the significant quadratic specifications. Notably, as Figure 5a shows, as large incumbents' R&D employment within 500 m increases from 55 (e^4) to 150 (e^5) workers, the estimated entry rate declines. Given that each incumbent in this size category contributes at least 51 R&D employees to the local environment, this implies that while a single large incumbent is a strong local attractor, additional large incumbents provide no further agglomerative impetus. Nevertheless, in contrast to H3, small incumbents' inventive activity is not an attractive force.

Finally, Table 5 compares the effects of inventive activity in the entrant's specialization by incumbents with and without university alliances. Here, the estimates support H3. For incumbents with a university alliance, only the linear term

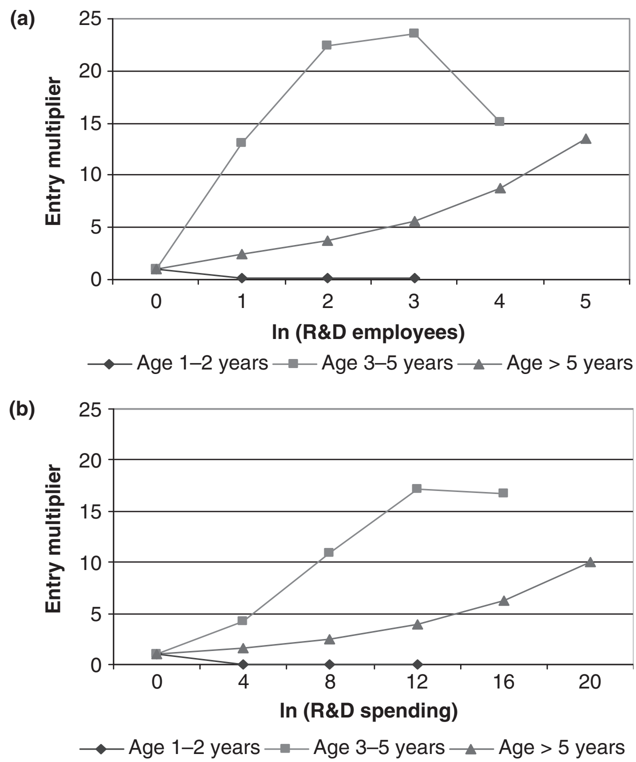

Figure 4 Multipliers for R&D employees within 500 m—controlling for age.

of the quadratic specification is significant, and the coefficient is positive. For incumbents without a university alliance, the coefficients for the linear terms are also significant and positive, but the squared terms are significant and negative, damping the positive effect.

Figures 6a and 6b depict these effects graphically, reporting multipliers estimated from Models 5a and 5b. The figures show the increasing agglomerative influence of inventive activity for incumbents with university alliances and the smaller and marginally declining effect of inventive activity by those without a university alliance. Thus, controlling for proximity to universities, the inventive activity of incumbents with university alliances has a larger agglomerative influence. This is consistent with the prediction in H3 that university alliances signal a local culture of learning and collaboration as well as potential access specialized knowledge and work with skilled scientists.

5. Discussion and conclusion

Although empirical research has demonstrated that concentrated regions of inventive activity attract greater entry of new firms, few studies have examined the patterns of

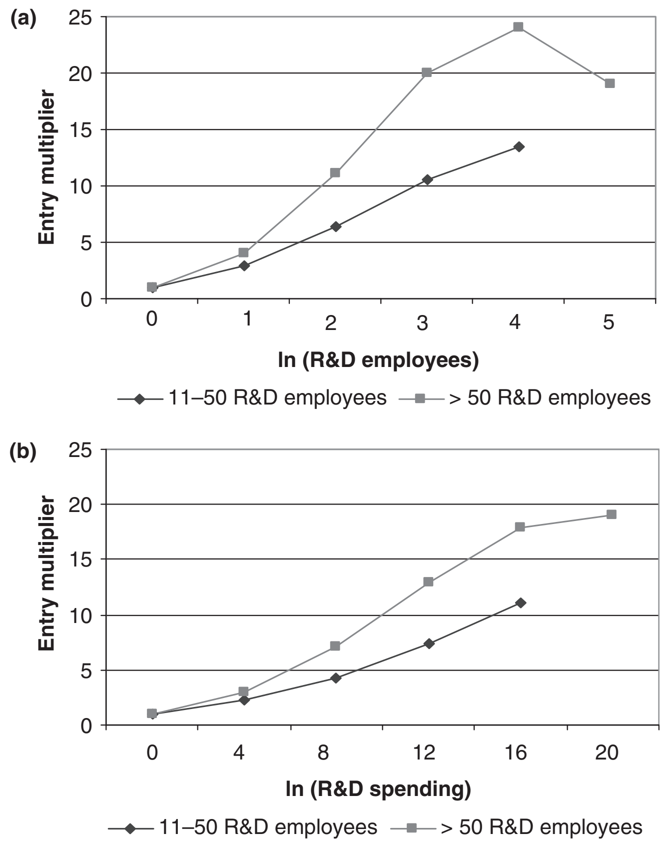

Figure 5 Multipliers for R&D employees within 500 m—controlling for size.

entry within those regions, or the role of potential spillovers in shaping entrants' choice of location. The logic of knowledge spillovers suggests that entrants will co-locate with active incumbents in their technological specialization. But knowledge spillovers are not unidirectional; entrants may be wary of neighbors capable of expropriating their knowledge. This tension brings into focus the strategic importance of location in knowledge-intensive, innovation-based industries.

We focused attention on these issues by using detailed data on Canadian biotechnology firms during the 1990s to examine the geographic scope and balance of spillover-seeking and expropriation-avoidance in entrants' locations. Our findings indicate that entrants in our sample strongly favored locations in close proximity—within 500 m—to inventively active incumbents in their specialization. And, moreover, that agglomeration effects dissipate rapidly, with inventive activity in the entrant's technological specialization beyond 500 m having little or no effect on entry. The results held whether inventive activity was measured as R&D employees

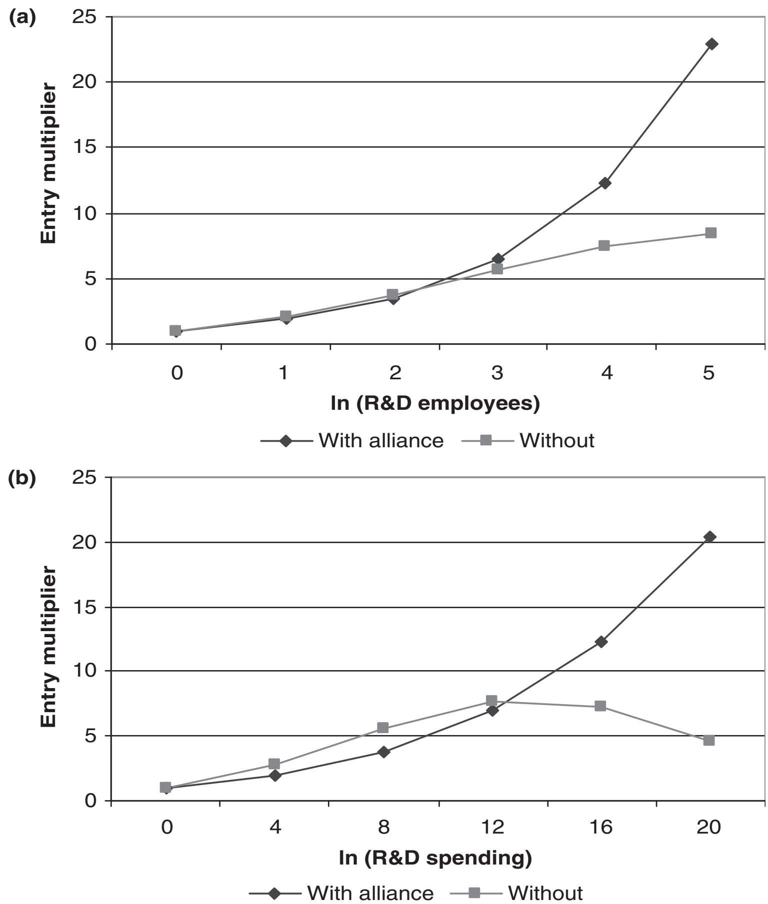

Figure 6 Multipliers for R&D employees within 500 m—controlling for university alliance.

or spending, but the magnitude of the effects was consistently larger, and somewhat more diffuse for R&D employees. The larger magnitude and slower decay R&D employee effects is consistent with the idea that knowledge spillovers flow through interactions among researchers within dense local social and professional networks (Saxenian, 1994; Liebeskind *et al.*, 1996; Owen-Smith and Powell, 2004).

Consistent with other recent micro-geographic studies, we find a rapid attenuation of attraction to a specific address. At this level of analysis, other sources of agglomeration such as labor market supply are invariant. The attraction is the presence of incumbents with a high potential to generate positive knowledge externalities. The potential for frequent face-to-face interaction, serendipitous encounters and easy scrutiny are facilitated by being in the neighborhood—near those firms working on similar things and open to sharing information. Notably, while the effects of incumbents' R&D activity attenuate rapidly, those of universities, while also attenuating rapidly, tend to diffuse more slowly, influencing entrants' locations up to 10 km away. Although it is possible that other processes are at work, our results support the view that knowledge spillovers requiring frequent face-to-face contact dissipate over short distances, while benefits of labor market pooling and shared inputs extend over a greater distance (Rosenthal and Strange, 2003). Thus, while Glaeser *et al.* (1992: 1127) observed that knowledge must cross corridors and streets more easily than ocean and continents, our results suggest a far tighter bounding. It is within small neighborhoods that the transmission of ideas we describe as knowledge spillovers appear to occur.

Extending prior research, our findings also indicate that two contextual factors increase entrants' tendency toward spillover-seeking at a neighborhood level, we argue, by mitigating expropriation concerns. One is increasing returns to positive information externalities that arise when incumbents in the same technological specialization concentrate geographically. Agglomerative effects of inventive activity in an entrant's specialization within 500 m were particularly strong when incumbents in the entrant's specialization were concentrated at the location. Knowledge spillovers depend on co-location of specialized firms sharing a relatively narrow common scientific or technological knowledge base. The greater value of potential knowledge spillovers within rather than across specialized technological applications creates advantages for firms located in areas of concentrated inventive activity in their technological area, leading to pronounced geographical clustering of technological specializations.

The other factor, for which support was mixed, is the entrepreneurial and open industrial organization that arises when inventively active younger and smaller incumbents and those with direct ties to universities concentrate geographically. Adolescent (3–5 years old) and larger (>50 R&D employees) incumbents, but not as predicted the youngest and smallest, had the greatest agglomerative influences. These estimates suggest more complex effects of incumbent demography on industrial organization than we had originally anticipated. Although it is plausible that incumbents must establish a track record of operations and achieve visibility before their inventive activity begins to affect agglomeration, such an argument still does not explain why it is the largest (and not mid-sized) incumbents' inventive activity that have the strongest agglomerative force. One possible explanation for this result is found in the recent agglomeration research on the so-called anchor tenant hypotheses (Feldman, 2003; Agrawal and Cockburn, 2003), which suggests that the presence of a large, local R&D-intensive incumbent may be important to both creating and capturing knowledge spillovers within particular locales.

More supportive of the industrial organization prediction, however, inventive activity was a stronger agglomerative force when incumbents in the entrant's specialization were allied with a university. Universities are a well-known agglomerative influence in high-technology fields and their role in facilitating knowledge transfer profound (e.g., Zucker *et al.*, 1998). Our findings reinforce the significance of university–industry alliances by pointing to their additional role in shaping a local industrial culture of learning and collaboration that signals a lower threat of expropriation by competitors and greater potential access to skilled university scientists. In sum, while empirical evidence of industrial organization effects of the kind recognized by Saxenian (1994) points to important sources of agglomeration overlooked in much prior work, our mixed results point to the need for further work in this area, for example, that considers the possibility of an optimal neighborhood mix, or balance of large and small and old and young incumbents, and that links Saxenian's ideas on industrial organization with recent work on anchor tenants.

Our findings have implications for the creation and development of industrial clusters that are relevant for both firms and for jurisdictions. Self-reinforcing virtuous circles that foster dense clustering appear to be created by the interaction of specialized R&D concentration, increasing returns and industrial organization. Increasing returns to existing concentrations of R&D activity and an open industrial organization encourage entrants to locate close by. Over time, the distribution of firms becomes more geographically dense, enhancing opportunities for interaction and the exchange of ideas, and further deepening the advantages of location. Entrants from one time period become the incumbents in the next time period and the neighborhood fills in and becomes more intensely focused and calibrated to that technological activity.

Many remote jurisdictions are investing resources to promote the formation of new biotechnology firms or to attempt to recruit firms from other locations. The insights our findings offer into factors shaping entrants' location choices may facilitate such efforts by enhancing entrants' willingness to co-locate. Concentrating resources around a single technological specialization may prove more useful than the dispersion of resources among a variety of technologies. But, the attractive force of technological concentration depends on the R&D activities and innovative capabilities of the firms; technological concentration is necessary, but not sufficient. Attracting a substantial cohort of adolescent firms or a single large anchor tenant, or facilitating university–industry alliances may also create an impetus. Again, however, the attractive force of industrial organization too is likely to be minimal without incumbents' investments to create a pool of knowledge. More generally, the Canadian economy may be too small to support a large number of diversified biotechnology clusters and so government policies aimed at fostering the creation of such clusters likely to be ill-advised, promoting a suboptimal geographic decentralization of firms. Support of specialized clusters may be warranted, however, particularly if coupled with support for development of related research capabilities at a nearby university. Overall, our findings point to the great challenges facing such jurisdictions; the most inventive firms are unlikely to be attracted to remote projects.

For firms, the implications of this research are twofold. First, our findings emphasize the importance of entry location in knowledge intensive, innovation-based industries. If location matters for innovation, it should matter most at the margin—for resource constrained entrants. Entrants may be able to enhance their choice of location relative to technologically similar incumbents by considering (i) their planned investments in R&D and technological focus and (ii) the potential for increasing returns and the development of informal social and more formal relationships with nearby firms. The strong localization of spillovers underscores the importance of local social networks and face-to-face communication in accessing agglomeration benefits. For incumbents, our results suggest that benefits from clustering can potentially be enhanced by demonstrating their commitment to R&D and by building a climate of trust and openness, for example through university–industry alliances, in order to attract new local entrants.

Several implications and directions for future research follow from our study, and also from its limitations. First, replications of our study in other settings and for other types of location choices would help to clarify the roles and establish the broader significance of increasing returns and industrial organization for the micro-dynamics of agglomeration. For example, as a neighborhood's concentration level or industrial organization change over time, do incumbents move in a manner consistent with our predictions for entry location? Additionally, while new entrants are initially unlikely to contribute more to knowledge spillovers than they can potentially absorb, as the balance of contributions to the pool of spillovers changes over time do entrants that increase their R&D intensity more than their neighbors tend to relocate?

In addition to such temporal shifts, clustering dynamics may also vary across specializations within an industry. One factor likely to materially affect patterns of co-location is the complexity of knowledge underlying industry innovations. While simple knowledge diffuses readily, more complex knowledge resists diffusion, sometimes even within the firm it originated. Knowledge of moderate complexity, however, diffuses unequally, following closely the lines of social and professional networks that facilitate the face-to-face interactions required for its effective transmission (Sorenson *et al.*, 2004). It may be that increasing returns are more important to facilitating clustering when knowledge is below moderate complexity, and so more diffuse spillovers likely to occur. In contrast, an open industrial organization may be more important to facilitating clustering when knowledge is of moderate complexity and above, making face-to-face channels essential to knowledge flows across organizational boundaries.

Differences in entrants' location choices also deserve further examination. It is likely that entrants differ in ways that affect their location choices (e.g., in their desire and capacity for interorganizational learning; in their location experience and preference). Consequently, the effects of proximity to incumbents' R&D investments and technological capabilities is unlikely to affect all entrants' location choices in the same way. Future work might thus usefully explore how entrant-specific features affect the propensity to seek spillovers. For example, as noted earlier, some entrepreneurship scholars suggest that entrants establish new ventures in areas where they happen to be located (e.g., Shane, 2002). Although our findings suggest that entrants may be more strategic about where they establish operations within a particular location, close physical proximity may be driven by more mundane factors unrelated to spillover seeking. Entrants may, for example, tend to locate new ventures close to previous employers or to spin off subsequent new ventures close by. Scientists may want to be able to commute easily between academic and corporate offices. Former students may locate new ventures close to their former universities. There may also be dedicated buildings receiving public subsidies to host new entrants. For example, the MaRS (Medical and Related Sciences) complex, which opened on the University of Toronto campus in 2005, houses specialized research and business incubation facilities with access to leading edge scientific equipment and co-located professional services, investors, technology transfer offices, and research and community networking organizations. Although some of these factors may be partly construed as spillover seeking, they may foster localized clustering that is not only or even primarily a result of spillover seeking. We accounted for such influences by controlling for entrants' proximity to universities as well as with region and entry group fixed effects. Future work explicitly designed to disentangle the effects of spillover seeking from such additional factors will contribute importantly to clarifying the role of local non-market knowledge externalities embraced by theories of industrial clustering in the new economic geography.

Also potentially informative in this regard may be comparative analyses of location patterns of *de novo* entrants and established firms' new plants or facilities, which vary in their R&D intensity and technological resources. Clustering under agglomeration economies is likely to disproportionately benefit certain firms. A firm that invests more in R&D compared to its competitors might disproportionately aid its competitors if the R&D investment spills over to its competitors' benefit. While new entrants are likely to benefit more from knowledge spillovers than they can potentially contribute, and so tend to seek spillovers, spillover-seeking may be less likely when established firms, which vary in their R&D intensity and technological resources, chose locations for new plants or facilities (Shaver and Flyer, 2000).

Our results indicate that Canadian biotechnology entrants in the 1990s strongly favored clustering. Did these entrants *desperately* seek spillovers? While strongly influenced by potential knowledge spillovers, our results suggest that entrants were influenced systematically by factors promoting the benefits of co-location, and sought locations that would allow them to benefit positively from knowledge spillovers. While our study reveals some new and more fine-grained dimensions of the location of new entrants, much work remains to be done. Our findings suggest that future studies be alert to both increasing returns produced by the concentration of specialized R&D activity and features of industrial organization that enhance knowledge externalities. We hope our analysis provides grist for those keen to better understand the micro-details of geographic clustering.

Addresses for correspondence

Barak S. Aharonson and Joel A. C. Baum, Rotman School of Management University of Toronto, 105 St. George Street, Toronto, ON M5S 3E6, CANADA. e-mail: barak.aharonson@rotman.utoronto.ca
Maryann P. Feldman, Institute of Higher Education, University of Georgia, Athens, GA 30602-6772, USA. e-mail: mfeldman@uga.edu

References

Agrawal, A. and I. Cockburn (2003), 'The anchor tenant hypothesis: examining the role of large, local, R&D-intensive firms in university knowledge transfer,' *International Journal of Industrial Organization*, **21**, 1227–1253.

Aldrich, H. E. and E. R. Auster (1986), 'Even dwarfs started small: liabilities of size and age and their strategic implications,' in B. M. Staw and L. L. Cummings (eds), *Research in Organizational Behavior*. Vol. 8, JAI Press: Greenwich, CT, pp. 165–198.

Almeida, P., G. Dokko and L. Rosenkopf (2003), 'Startup size and the mechanisms of external learning: increasing opportunity and decreasing ability?' *Research Policy*, **32**, 301–315.

Almeida, P. and L. Rosenkopf (2003), 'Overcoming local search through alliances and mobility,' *Management Science*, **49**, 751–766.

Audretsch, D. B. and M. P. Feldman (1996), 'R&D spillovers and the geography of innovation and production,' *American Economic Review*, **86**, 630–640.

Arzaghi, M. and J. V. Henderson (2004), 'Networking off Madison Avenue,' Working Paper. Brown University.

Baum, J. A. C. and H. A. Haveman (1997), 'Love thy neighbor? differentiation and agglomeration in the Manhattan hotel industry, 1898–1990,' *Administrative Science Quarterly*, **42**, 304–338.

Baum, J. A. C. and C. Oliver (1991), 'Institutional linkages and organizational mortality,' *Administrative Science Quarterly*, **36**, 187–218.

Baptista, R. and P. Swann (1998), 'Do firms in clusters innovate more?' *Research Policy*, **27**, 525–540.

Baptista, R. and P. Swann (1999), 'The dynamics of firm growth and entry in industrial clusters: a comparison of the US and UK computer industries,' *Journal of Evolutionary Economics*, **9**, 73–399.

Beaudry, C. (2001), 'Entry, growth and patenting in industrial clusters: A study of the aerospace industry in the UK,' *International Journal of the Economics of Business*, **8**, 405–436.

Beaudry, C. and S. Breschi (2003), 'Are firms in clusters really more innovative?' *Economics of Innovation and New Technology*, **12**, 325–342.

Breschi, S. and F. Lissoni (2001), 'Knowledge spillovers and local innovation systems: A critical survey,' *Industrial and Corporate Change*, **10**, 975–1005.

Brown, J. S. and P. Duguid (2000), *The Social Life of Information*. Harvard Business School Press: Cambridge, MA.

Buenstorf G. and S. Klepper (2005), 'Heritage and agglomeration: The Akron Tire Cluster Revisited,' Papers on Economics and Evolution, Max Planck Institute of Economics, Evolutionary Economics Group.

Chen, M. J. and D. Hambrick (1995), 'Speed, stealth and selective attack: how small firms differ from large firms in competitive behavior,' *Academy of Management Journal*, **38**, 453–482.

Chinitz, B. (1961), 'Contrasts in agglomeration: New York and Pittsburgh,' *American Economic Review*, **51**, 279–289.

Cohen, W. M. and D. A. Levinthal (1990), 'Absorptive capacity: a new perspective on learning and innovation,' *Administrative Science Quarterly*, **35**, 128–152.

Delmar, F. and S. Shane (2003), 'Does business planning facilitate the development of new ventures?' *Strategic Management Journal*, **24**, 1165–1185.

Dumais, G., G. Ellisson and E. L. Glasear (2002), 'Geographic concentration as a dynamic process,' *Review of Economics and Statistics*, **84**, 193–204.

Duranton, G. and H. G. Overman (2005), 'Testing for localization using micro-geographic data,' *Review of Economic Studies*, **72**, 1077–1106.

Feldman, M. P. (1994), *The Geography of Innovation*. Kluwer Academic Publishers: Dordrecht.

Feldman, M. P. (2003), 'The locational dynamics of the U.S. biotech industry: knowledge externalities and the anchor hypothesis,' *Industry and Innovation*, **10**, 311–328.

Feldman, M. P. and D. B. Audretsch (1998), 'Innovation in cities: science-based diversity, specialization, and localized competition,' *European Economic Review*, **43**, 409–429.

Flyer, F. and J. M. Shaver (2003), 'Location choices under agglomeration economies and strategic interaction,' in J. A. C. Baum and H. R. Greve (eds), *Geography and Strategy: Advances in Strategic Management*. Vol. 20, Elsevier/JAI Press: Oxford, UK, pp. 175–195.

Gatrell, A. C., T. C. Bailey, P. J. Diggle and B.S. Rowlingson (1996), 'Spatial point pattern analysis and its application in geographical epidemiology,' *Transactions of the Institute of British Geographers*, **21**, 256–274.

Glaeser, E. L., H. D. Kallal, J. A. Scheinkman and A. Shleifer (1992), 'Growth in cities,' *Journal of Political Economy*, **100**, 1126–1152.

Hannan, M. T. and J. Freeman (1977), 'The population ecology of organizations,' *American Journal of Sociology*, **82**, 929–964.

Hannan, M. T. and J. Freeman (1984), 'Structural inertia and organizational change,' *American Sociological Review*, **49**, 149–164.

Henderson, J. V. (1986), 'Efficiency of resource usage and city size,' *Journal of Urban Economics*, **19**, 47–70.

Jacobs, J. (1969), *The Economy of Cities*. Vintage: New York.

Jaffe, A. (1989), 'Real effects of academic research,' *American Economic Review*, **79**, 957–970.

Jensen, M. (2003), 'The role of network resources in market entry: commercial banks' entry into investment banking, 1991–1997,' *Administrative Science Quarterly*, **48**, 466–497.

Kalnins, A. and W. Chung (2004), 'Resource-seeking agglomeration: a study of market entry in the lodging industry,' *Strategic Management Journal*, **25**, 689–699.

King, G. and L. Zeng (2001), 'Logistic regression in rare events data,' *Political Analysis*, **9**, 1–27.

Kogut, B. (2000), 'The network as knowledge: generative rules and the emergence of structure,' *Strategic Management Journal*, **21**, 405–425.

Krugman, P. (1991), 'Increasing returns and economic geography,' *The Journal of Political Economy*, **99**, 483–499.

Lancaster, T. and G. Imbens (1996), 'Efficient estimation and stratified sampling,' *Journal of Econometrics*, **74**, 289–318.

Liebeskind, J. P., A. L. Oliver, L. Zucker and M. Brewer (1996), 'Social networks, learning, and flexibility: sourcing scientific knowledge in new biotechnology firms,' *Organization Science*, **7**, 428–443.

Lucas Jr, R. (1988), 'On the mechanics of economic development,' *Journal of Monetary Economics*, **22**, 2–42.

Marshall, A. (1920), *Principles of Economics*. Macmillan: London.

Mowery, D. C. and A. A. Ziedonis (2001), The Geographic reach of market and Non-Market Channels of Technology Transfer: Comparing Citations and Licenses of University Patents. NBER working paper No. 8568.

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

Porter, M. (1990), *The Competitive Advantage of Nations*. Macmillan: London.

Porter, M. E. and S. Stern (2001), 'Innovation: location matters,' *Sloan Management Review*, **Summer**, 28–36.

Ring, P. S. and G. P. Rands (1989), 'Sensemaking, understanding, and committing: emergent interpersonal transaction processes in the evolution of 3M's microgravity research program,' in A. H. Van de Ven, H. L. Angle and M. S. Poole (eds), *Research on the Management of Innovation: The Minnesota Studies*. Harper & Row: New York, pp. 337–366.

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

Romer, P. (1990), 'Endogenous technological change,' *Journal of Political Economy*, **98**, S71–S102.

Rosenthal, S. S. and W. C. Strange (2003), 'Geography, industrial organization, and agglomeration,' *Review of Economics and Statistics*, **85**, 377–393.

Saxenian, A. (1994), *Regional Advantage*. Harvard University Press: Cambridge, MA.

Shane, S. (ed) (2002), *Foundations of Entrepreneurship*. Edward Elgar: Aldershot, UK.

Shaver, J. M. and F. Flyer (2000), 'Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States,' *Strategic Management Journal*, **21**, 1175–1193.

Sorenson, O. (2003), 'Social networks and industrial geography,' *Journal of Evolutionary Economics*, **13**, 513–527.

Sorenson, O., J. W. Rivkin and L. Fleming (2004), 'Complexity, networks and knowledge flow.' Working paper. Anderson School of Management, UCLA and Harvard Business School.

Stinchcombe, A. L. (1965), 'Social structure and organization,' in J. G. March (ed.), *Handbook of Organizations*. Rand McNally: Chicago, pp. 142–193.

Stuart, T. E. and O. Sorenson (2003), 'The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms,' *Research Policy*, **32**, 229–253.

Surico, P. (2003), 'Geographic concentrations and increasing returns,' *Journal of Economic Surveys*, **17**, 693–709.

Tomz, M. (2003), 'ReLogit: rare events logistic regression,' *Journal of Statistical Software*, **8**(2), 137–163.

Wallsten, S. J. (2001), 'An empirical test of geographic knowledge spillovers using geographic information systems and firm-level data,' *Regional Science and Urban Economics*, **31**, 571–599.

Wonnacott, R. T. and T. H. Wonnacott (1970), *Econometrics*. 2nd edn. John Wiley & Sons: New York.

Yoffie, D. (1993), 'Foreign direct investment in semiconductors,' in Kenneth Froot (ed.), *Foreign Direct Investment*. University of Chicago Press: Chicago.

Zucker, L. G. (1986), 'Production of trust: institutional sources of economic structure, 1840 to 1920,' in B. Staw and L.L. Cummings (eds), *Research in Organizational Behavior*. Vol. 8, JAI Press: Greenwich, CT, pp. 55–111.

Zucker, L. G., M. Darby and M. Brewer (1998), 'Intellectual human capital and the birth of US biotechnology enterprises,' *American Economic Review*, **88**, 290–306.

Appendix

Table A1 Descriptive statistics—R&D employee variables

		Entry postal codes		Random non-entry postal codes		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		Mean	S.D.	Mean	S.D.																		
Inventive activity—entrant's specialization																							
1	<500 m	0.911	2.548	0.178	0.802	1.00																	
2	500 m to 2 km	1.154	2.624	1.176	2.529	0.09	1.00																
3	2–10 km	6.507	9.062	7.497	9.812	0.07	0.19	1.00															
4	<500 m (Age 1–2 Years)	0.228	0.792	0.090	0.477	0.65	0.05	0.09	1.00														
5	<500 m (Age 3–5 Years)	0.321	1.021	0.048	0.410	0.73	0.07	0.03	0.25	1.00													
6	<500 m (Age >5 Years)	0.362	1.445	0.040	0.364	0.77	0.07	0.05	0.22	0.35	1.00												
7	<500 m (1–10 R&D employees)	0.213	0.706	0.085	0.413	0.68	0.09	0.10	0.70	0.40	0.40	1.00											
8	<500 m (11–50 R&D employees)	0.419	1.578	0.051	0.421	0.81	0.04	0.02	0.54	0.53	0.66	0.40	1.00										
9	<500 m (>50 R&D employees)	0.279	1.001	0.042	0.448	0.61	0.07	0.05	0.16	0.61	0.53	0.17	0.17	1.00									
10	<500 m (University alliance)	0.455	1.399	0.087	0.495	0.83	−0.01	0.11	0.60	0.65	0.56	0.53	0.80	0.38	1.00								
11	<500 m (No university alliance)	0.457	1.454	0.091	0.563	0.86	0.15	0.02	0.50	0.59	0.74	0.62	0.58	0.65	0.43	1.00							
Concentration—entrant's specialization																							
12	<500 m	0.006	0.018	0.001	0.007	0.75	0.05	0.02	0.49	0.49	0.62	0.58	0.64	0.35	0.62	0.64	1.00						

(continued)

Table A1 Continued

		Entry postal codes		Random non-entry postal codes		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		Mean	S.D.	Mean	S.D.																		
Inventive Activity—other specializations																							
13	<500 m	2.823	4.167	1.165	2.973	0.22	0.22	−0.01	0.31	0.08	0.10	0.30	0.13	0.07	0.22	−0.01	0.15	1.00					
14	500 m to 2 km	8.429	11.012	7.256	10.505	0.01	0.36	0.02	0.10	−0.03	−0.03	0.09	−0.03	−0.01	0.43	0.00	0.07	0.30	1.00				
15	2–10 km	26.595	23.031	28.372	25.578	−0.05	0.06	0.27	−0.05	−0.04	−0.02	−0.03	−0.05	−0.03	0.06	0.26	−0.01	0.08	0.39	1.00			
Number of universities																							
16	<500 m	0.155	0.363	0.034	0.202	0.11	0.10	−0.09	0.05	0.16	0.04	0.12	0.07	0.06	0.10	−0.09	0.10	0.09	0.14	0.00	1.00		
17	500 m to 2 km	0.481	1.076	0.479	1.032	0.06	0.33	0.04	0.14	0.02	0.00	0.14	0.04	−0.02	0.33	0.04	0.09	0.44	0.50	0.13	0.11	1.00	
18	2–10 km	0.893	1.343	1.031	1.433	−0.05	0.06	0.35	−0.04	−0.04	−0.04	−0.05	−0.05	−0.01	0.06	0.35	−0.05	−0.10	0.25	0.51	−0.09	−0.01	1.00

Note: The sample included 1236 observations (206 entry postal codes; 1030 random non-entry postal codes).

Table A2 Descriptive statistics—R&D spending variables

		Entry postal codes		Random non-entry postal codes		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
		Mean	S.D.	Mean	S.D.																		
Inventive activity—entrant's specialization																							
1	<500 m	4.381	11.820	0.838	3.921	1.00																	
2	500 m to 2 km	4.657	10.642	5.173	11.173	0.06	1.00																
3	2–10 km	28.718	41.944	32.828	44.648	0.12	0.21	1.00															
4	<500 m (Age 1–2 Years)	1.053	4.130	0.406	2.310	0.63	0.01	0.15	1.00														
5	<500 m (Age 3–5 Years)	1.492	4.771	0.266	2.093	0.65	0.03	0.04	0.18	1.00													
6	<500 m (Age >5 Years)	1.836	7.567	0.165	1.747	0.75	0.08	0.07	0.19	0.20	1.00												
7	<500 m (1–10 R&D employees)	1.072	3.840	0.513	2.689	0.68	0.05	0.15	0.71	0.39	0.33	1.00											
8	<500 m (11–50 R&D employees)	2.116	7.757	0.201	2.107	0.78	0.03	0.02	0.44	0.44	0.68	0.23	1.00										
9	<500 m (>50 R&D employees)	1.193	4.279	0.124	1.494	0.54	0.05	0.11	0.07	0.52	0.48	0.16	0.17	1.00									
10	<500 m (University alliance)	2.472	7.603	0.598	3.174	0.87	0.01	0.15	0.69	0.56	0.54	0.58	0.72	0.41	1.00								
11	<500 m (No university alliance)	1.908	6.206	0.240	2.021	0.75	0.10	0.04	0.28	0.48	0.71	0.51	0.53	0.49	0.33	1.00							
Concentration—entrant's specialization																							
12	<500 m	0.006	0.018	0.001	0.007	0.72	0.05	0.02	0.38	0.45	0.61	0.45	0.58	0.39	0.61	0.55	1.00						

(continued)

Table A2 Continued

	Entry postal codes		Random non-entry postal codes		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
	Mean	S.D.	Mean	S.D.																		
Inventive activity—other specializations																						
13 <500 m	11.580	18.271	3.873	10.722	0.26	0.11	−0.04	0.28	0.09	0.17	0.29	0.14	0.10	0.11	−0.04	0.16	1.00					
14 500 m to 2 km	34.597	44.051	29.496	42.611	0.00	0.31	0.01	0.03	0.01	−0.03	0.06	−0.04	0.00	0.35	−0.05	0.07	0.24	1.00				
15 2–10 km	117.404	102.906	124.238	111.493	−0.06	0.06	0.21	−0.07	−0.04	−0.02	−0.04	−0.05	−0.03	0.07	0.21	−0.01	0.07	0.39	1.00			
Number of universities																						
16 <500 m	0.155	0.363	0.034	0.202	0.10	0.09	−0.10	0.06	0.16	0.01	0.12	0.04	0.06	0.09	−0.10	0.10	0.12	0.12	0.00	1.00		
17 500 m to 2 km	0.481	1.076	0.479	1.032	0.03	0.25	0.01	0.06	0.02	−0.02	0.08	−0.01	−0.01	0.25	0.01	0.09	0.22	0.37	0.13	0.11	1.00	
18 2–10 km	0.893	1.343	1.031	1.433	−0.05	0.06	0.34	−0.05	−0.02	−0.03	−0.05	−0.04	0.00	0.06	0.34	−0.05	−0.11	0.19	0.47	−0.09	−0.01	1.00

Note: The sample included 1236 observations (206 entry postal codes; 1030 random non-entry postal codes).