1-1-2000

The Evolution of Hot Spots in the U.S. Biotechnology Industry

Timothy B. Folta
Purdue University

Yoon-Suk Baik
Purdue University

The Evolution of Hot Spots and Blind Spots in the U.S. Biotechnology Industry

Timothy B. Folta
Krannert Graduate School of Management
Purdue University
West Lafayette, IN 47907-1310
765-494-9658
folta@mgmt.purdue.edu

Yoon-Suk Baik
Krannert Graduate School of Management
Purdue University
West Lafayette, IN 47907-1310
765-496-7723
baik@purdue.edu

July 27, 2001

Abstract

This study empirically tests Pouder & St. John's propositions (1996) on the evolutions of geographic clusters, in the context of the U.S. biotechnology industry. We find that during the origination period of the cluster evolution clustered biotechnology firms exhibit higher cost economies and legitimacy in obtaining resources than non-clustered competitors. While, after the early period, such clustered firms' advantage in resource access and innovations are statistically significantly declining. These findings evidence the decline of positive net benefits from geographic clustering over time. Given the inconsistent empirical results in the literature on whether there are positive net benefits to geographically clustered firms, this study sheds light on the importance of time dimension in geographic clustering to understand the net benefits of geographic clustering. Since economies and diseconomies of agglomeration change over time, the net benefits of geographic clustering can be time-variant, possibly leading to different empirical results if the evolution of geographic clustering is not appropriately considered. Accordingly, this study answers a call to empirical studies on dynamics of geographic clustering.

In spite of increased attention on geographic concentration in industries, there is little consensus on the benefits of geographic clustering to clustered firms. Recently, Staber (1998) and Shaver & Flyer (2000) find higher organizational failure rates within geographic clusters in the context of the German knitwear industry and the foreign direct investments in the U.S. manufacturing industries, respectively. These results are seemingly inconsistent with empirical findings of benefits from clusters in early studies (e.g., Hill & Naroff, 1984; Jaffe, Trajtenberg & Henderson, 1993; Audretsch & Feldman, 1996, DeCarolis & Deeds, 1999). This empirical inconsistency raises the question, whether geographic clustering is really beneficial to clustered firms and regions. By focusing on the evolution of geographic clusters over time, this study attempts to suggest a way to reconcile the inconsistent findings of the net benefits from geographic clustering.

Given very few theoretical studies on the evolution of geographic clusters, we consider that the theory proposed by Pouder & St. John (1996) is one of the earliest and compelling attempts to address the time-variant benefits of geographic clustering. They argue that the benefits to geographic clustering vary by stages of development of the clusters. During the early stage of geographic clusters (i.e., the origination phase), economies of agglomeration (or, benefits of clusters) dominate diseconomies of agglomeration (or, costs of clusters), showing positive net benefits of clusters. As time goes by, clustered firms are likely to experience relatively high levels of congestion costs, within-region competition for localized inputs, and knowledge expropriation, expecting decline in the net benefits from geographic clustering in the later phases (i.e., the convergence and the decline phases). Accordingly, we consider that geographic clustering can be either advantageous or disadvantageous to clustered firms, depending on the phases of the evolutionary path of geographic clustering.

In this study, we empirically test relevant hypotheses derived from their theory, in the context of the biotechnology industry. First, our attempt to empirically investigate the evolutionary path of geographic clustering can verify the importance of time dimension of the benefits of clustering and can suggest an approach to understand seemingly inconsistent findings on the net benefits from geographic clustering.

Second, this study sheds light on empirical studies to test the effect of the cluster evolution. Most of existing empirical works have concentrated on events at one point in time to evaluate the effects of geographic clusters, such as clustered and non-clustered firms' performance in initial public offerings (e.g., DeCarolis & Deeds, 1999) or their long-term survival (e.g., Shaver & Flyer, 2000). These approaches, however, can neglect the economic consequences of geographic clustering in the course of its evolutionary process. If the benefits of clusters change during a considerable length of period, then the static approaches cannot capture the economically significant time-varying characteristics of clusters. Lastly, we also believe that our study can be significant to management practices and public policy arena, for better evaluation of the net agglomeration economies.

In what follows, we discuss theory and hypotheses on the changes in the net benefits of geographic clustering, based on the theory of Pouder & St. John (1996). Empirical methods including data and empirical designs follow in the next section. Empirical results are summarized in the following section. We conclude the study with discussion about limitations and issues for future studies.

Theory and Hypotheses

As suggested by Marshall (1920) and Krugman (1994), among others, geographic clustering is an increasing function of the net benefits to competing firms within a region. If positive net benefits are expected from a geographic cluster, new entry will occur within the cluster, enhancing geographic clustering. The net benefits to clustering are determined by the benefits (economies) and the costs (diseconomies) of agglomeration. The main sources of agglomeration economies are pooled labor forces, specialized suppliers, and knowledge spillovers within a cluster (Krugman, 1994; Prevezer, 1997). Firms can obtain quality labor and other input factors at lower costs within a clustered region (Porter, 1998). The proliferation of innovations within the region can be possible through localized knowledge spillovers via labor shifts (Almedia & Kogut, 1999), interaction with research institutions (Audretsch & Feldman, 1996), or interfirm relationship (Scott, 1989) within the region. Whereas, diseconomies of agglomeration are generated by congestion costs, increased competition, and knowledge expropriation (Prevezer, 1997; Shaver & Flyer, 2000). As competing firms gather within a region, competition for input factors will increase. The risks of knowledge expropriation by adjacent competitors also emerge with increased knowledge spillovers.

Given the theory of agglomeration, Pouder & St. John (1996) suggest that the strength of economies and diseconomies of agglomeration varies with the age of geographic clusters. They propose three distinct phases of the evolution of geographic clusters: origination, convergence, and decline. During the origination phase, a cluster exhibits high growth rate and innovative activities within the region, because a firm's fundamental processes in resource access, legitimacy, and strategy formulation within the cluster show cost economies and more innovative strategy formulation.

Over time, the growth rate and innovative activities of the cluster stabilize in the convergence phase, as costs increase within the cluster due to congestion in resource access, imitative behaviors from isomorphism, and inertia in strategic formulation. This convergence phase ultimately leads to the decline of clusters that limit clustered firms' resource-access, legitimating, and innovative capabilities and adversely affect the growth of clusters. Accordingly, we consider that the growth of geographic clusters reflect the net benefits to geographic clustering that will vary across phases of cluster development. Then, we can expect the following hypothesis.

Hypothesis 1: *The growth rate of geographic clustering in an industry is declining over time.*

Regarding the time-varying patterns of the growth of geographic clusters, we need to specify the characteristics of each phase of the cluster evolution over time. We consider that Pourder & ST. John's (1996) theory on the evolutionary path of clusters is worthwhile to elaborate. As discussed above, the characterization of the cluster evolution is critical for better estimation of the effects of geographic clustering. A study focusing on the cluster effects with respect to long-term firm performance (e.g., Shaver & Flyer, 2000) can only predict the ultimate effects of the geographic clustering doomed to be deteriorated. This approach, however, cannot capture the positive economic consequences of geographic clustering in the course of the cluster development. Whereas, a study without time-varying notion of the effects of clusters (e.g., DeCarolis & Deeds, 1999) can also be disadvantageous in terms of its ignorance of time dimension of geographic clustering. In this regard, we set out to test Pouder & St. John's major propositions on the phases of the cluster evolution. Since their theory mainly focuses on the origination and the convergence phase of the cluster development, we concentrate on the first two phases.

Origination phase: According to Pouder & St. John (1996), the origination phase begins with success by the initial firm(s) that can induce qualified suppliers, skilled labors, and informed investors. This lowers the cost of entry for subsequent firms. In addition, firms locating in the cluster can enhance legitimacy through relationship with firms within the region. Clustered firms can share regional ties to a research base (e.g., research universities), a skilled labor pool, a network of qualified suppliers, and an informed group of venture capitalists. Lastly, more informed strategy formulation encourages the emergence of the cluster. During the origination phase, more information within the cluster will help better strategy formulation through mobile labor force, social interaction, cooperative alliances, direct observation, and local media.

To the contrary, competitors outside the cluster will face high costs for hiring specialized employees and for transacting with suppliers and researchers, during the origination phase. Firms outside the cluster also find difficulties in imitating the complex routines involved in the infrastructure of the clustered firms and have more imperfect information than clustered firms when identifying specialized labors and qualified suppliers and innovation opportunities, struggling to maintain competitive parity with the competitors in the cluster. Thus, in the cluster, more entry will be likely due to large benefits from the cluster than outside the cluster. Accordingly, we consider following hypotheses originated from Pouder & St. John (1996).[1]

Hypothesis 2: *During the origination phase of the cluster evolution, geographically clustered firms within the cluster will experience greater cost economies and legitimacy than industry competitors that are outside of the cluster.*

Hypothesis 3: *During the origination phase, the rate of growth in number of competitors within the cluster will exceed the rate of growth in numbers of competitors outside the cluster.*

Hypothesis 4: *During the origination phase, clustered firms will be responsible for an increasing proportion of industry innovations, compared to non-clustered competitors.*

[1] Following hypotheses are basically duplicated from the propositions by Pouder & St. John (1996) for empirical tests.

Convergence phase: As the clusters move through the convergence phase, advantages of clusters dissipate largely in three ways. First, the benefits from low costs to access resources within the clusters will erode because congestion in the cluster bound the cost economies. Second, legitimating processes also lead to isomorphism within the clusters and hence induce imitative rather than innovative behaviors. Lastly, strategy formulation within the clusters becomes homogenous and biased toward the strategies of the clustered firms. This leads clustered firms to be less flexible in adjusting to environmental changes. Firms outside the cluster are independent of this kind of adverse selection process within regions (Shaver & Flyer, 2000).

Given the diminished benefits of the clusters, high density of existing firms in the clusters after the origination phase intensifies localized competition in the clusters. As knowledge spills over across regions, competitors outside the clusters will relatively recover from strategic disparity. Thus, the pattern of entry and exit in the cluster and in the overall industry will reach parity, and there will be no overall advantages for the clusters. That is, the patterns of growth for both the clusters and the industry overall will become similar. Consequently, we can expect following hypotheses.[2]

Hypothesis 5: *During the convergence phase, the agglomeration economies in the cluster will erode, and the hot spot firms will experience cost economies similar to competitors outside of clusters.*

Hypothesis 6: *During the convergence phase, cluster growth rate will stabilize compared to the larger industry population.*

[2] As above, following hypotheses are basically duplicated from the propositions by Pouder & St. John (1996) for empirical tests.

Hypothesis 7: *During the convergence phase, the collective rate of innovation emanating from clusters will decrease over time.*

Methods

Industry

We test above seven hypotheses in the context of the biotechnology industry. This industry is appropriate for our empirical design. First, the industry is geographically concentrated as over ninety percent of the population is clustered in nineteen major and minor clusters (Burrill & Lee, 1993). Many studies have identified and examined geographic clustering in this industry (e.g., Audretsch & Stephan, 1996; DeCarolis & Deeds, 1999; Zucker, et al., 1999). In addition, the dominance of small firms in the industry (seventy five percent of small firms in the total population (Burrill & Lee, 1993)) provides a relevant empirical context to test the theory proposed by Pouder & St. John (1996). In their theory, one of the driving forces to time-varying net benefits of geographic clusters consists of cost economies in resource procurement and obtaining legitimacy. As pointed by Stuart, Huang, & Hybels (1999), small biotechnology ventures are very sensitive to resource access and legitimacy. In this regard, the industry can be an appropriate domain to study changes in the net benefits of clusters over time.

The patterns of the evolution of geographic clusters reflect important events in the industry history. The contemporary biotechnology industry started with two radical innovations – recombinant DNA (r-DNA) and cell fusion. In 1973, Boyer and Cohen introduced r-DNA, genetic material from one cell into the DNA structure of another. As such, this event is assumed to be the beginning of the industry (Stuart, et al., 1999).

Milstein and Kohler succeeded in a second-generation technique of cell fusion that creates a hybrid cell capable of producing highly purified proteins in two years later. The year of 1980, however, induced true spurt in the number of new founding firms in the industry, after 1) the U.S. Supreme Court decision that a new life form can be patentable, 2) the passage of the Patent and Trademark Amendment Act of 1980 that enabled universities to apply for patents, and 3) Genentech's successful initial public offering. A major third-generation technique of protein engineering encouraged subsequent entry during 1980s. Investors of the stock markets, however, constrained the speed of the industry growth during the 1990s.[3]

Sample

The data describe 825 U.S. biotechnology firms founded between 1973 and 1997. We include both private and public firms founded from industry origin to most recent years. Our data incorporate all biotechnology firms regardless of their market segments, such as therapeutic, diagnostic, agricultural, veterinary, food-process, and others. This allows a comprehensive and representative sample from the industry population, to study the evolution of geographic clustering in the industry. The data set is established, mostly based on Bioscan directory (published by Oryx Press) and the Actions database (published by the North Carolina Biotechnology Center (NCBC)). These data sources are commonly used for studying the biotechnology industry (e.g., Stuart, et al., 1999; Zucker, et al., 1999).

[3] For further review of the biotechnology industry history, see Stuart, et al. (1999) and Ryan, Freeman, & Hybels (1995).

Analysis

Geographic clusters. Our first task is to identify geographic clusters for analysis. There is no agreement on how to cluster firms geographically. In the literature, most studies have relied on the state as the unit of location (e.g., Krugman, 1994; Audretsch & Feldman, 1996; Prevezer, 1997). Others have used the Metropolitan Statistical Area (e.g., DeCarolis & Deeds, 1999) or Economic Functional Area defined by the Department of Commerce (e.g., Zucker, et al., 1999). These proxies for identifying geographic clusters can sometimes be inadequate, if agglomeration effects are not constrained by such conventionally identified boundaries. We consider that agglomeration economies are centered on spatial proximity. Accordingly, we cluster firms by identifying geographic distance based on each sample firm's zip code.

In this study, we mostly use the first two digits of zip codes to identify geographic clusters. In case that many states are located within relatively adjacent areas (e.g., New England and Mid-Atlantic areas), we use three digits of zip codes that capture more plausible distances related to agglomeration economies.[4] As shown in Table 1, this process identifies nineteen major and minor clusters that accommodate at least ten biotechnology firms within the regions. These identified clusters are similar to those identified by Burrill & Lee (1991, 1992, & 1993) and Audretsch & Stephan (1996).[5] In particular, ten largest clusters that we identified exactly overlap with top ten clusters in the survey by Burrill & Lee (1993).

[4] The first two digits of five-digit zip codes usually divide a state into two regions, according to the U.S. Postal Service. The third digit in five-digit zip code stands for a post office that govern postal services within a certain, equally divided region affiliated to a two-digit sub-region.

[5] Biotechnology firms are geographically concentrated in three primary regions (the San Francisco Bay Area, San Diego and Boston), two secondary regions (Philadelphia and New York), and a number of small clusters (Audretsch & Stephan, 1996). Ernst & Young identify three primary regions, two secondary regions, several other regions with at least 20 companies, and a host of small clusters (Burrill & Lee, 1992).

Since Pouder & St. John (1996) implicitly dichotomize an industry into clustered and non-clustered firms, we need to operationalize the concept of a geographic cluster and clustered firms. Following Burill & Lee's (1993: vi) assumption, we use a criterion that a regional unit is a cluster if it accommodates more than 20 competing firms. Similarly, firms in clusters become clustered firms, while otherwise firms are non-clustered firms. From this criterion, the ten largest clusters in our sample are regarded as clusters. As DeCarolis & Deeds (1999) did, we treat these ten largest clusters as main geographic clusters that show a significant level of agglomeration economies. The other nine minor clusters and other regions are assumed to be non-clustered regions. [6]

[Insert Table 1 about here]

Identification of phases. To identify different phases of the evolution of clusters, we need empirical definition of the phase of origination, convergence, and decline. According to the theory of Pouder & St. John (1996:1196), each phase of the cluster development is defined by the difference in growth between clustered and non-clustered firms over time. If the difference in growth between clustered and non-clustered firms increases, clusters are in the origination phase. If the difference in growth between them starts to decrease, clusters begin to experience the convergence phase up to the point at which the level of growth between clustered and non-clustered firms are equal. Beyond that point, the dominant growth of non-clustered firms against clustered firms signifies the decline phase. Figure 1 graphically describes these operational definitions.

[6] For robustness, we considered another case that the six largest clusters are assumed to be clusters while other regions are non-clustered. We had similar results while not reported. Whereas, if we consider nineteen major and minor regions clusters (based on a criterion that a regional unit having more than ten firms is assumed to be a cluster), we had significantly different results unreported in this study. We conjecture that significant level of agglomeration economies comes from a region clustered with more than ten competing firms in our sample.

[Insert Figure 1 about here]

In this study, we identify the ending year of the origination phase by demarking the year in which the maximum difference in growth between clustered and non-clustered firms is observed during the study period. We measure the growth of clustered and non-clustered firms by number of firms in a cluster and non-clustered regions, as in DeCarolis & Deeds (2000), among others. Since we have ten different clusters, we figure out the maximum difference for each of ten clusters, by subtracting the average number of firms in non-clustered regions - cluster 11 through 19 and non-clustered regions - from the number of firms in a cluster. Then, we obtain the ending year of the origination phase for each of ten clusters. The ending years of the origination period are shown in Table 1.

If the ending year of the origination phase coincides with the ending year of our study period, clusters in the industry are still in the origination phase. If the ending year of the origination phase falls into a year within our study period, clusters are in the convergence phase after passing through the origination phase. Our data exhibits no such points that the difference of growth between clustered and non-clustered firms becomes zero (i.e., decline phase). Accordingly, during our study period, the biotechnology industry appears to enter the convergence phases after the origination phase. Detailed features of the cluster evolution in the industry are shown in Table 1.

Growth of clusters. To test hypothesis 1 on the overall growth pattern of clustered firms over time, we measure the growth of clusters by the number of existing biotechnology firms within clusters. Organizational ecologists commonly use the number of existing firms in a cluster to measure the density of firms in a region (Carroll & Hannan, 1995; Lomi, 1995). The underlying assumption of this measure is that firms are relatively homogenous and no scale economies exist in the industry. If firms are significantly different in terms of their size and scale economies are important in the industry, the number of existing firms can mislead the growth of clusters. The agglomeration of small number of large firms may imply higher growth of a cluster than the agglomeration of large number of small firms within a region. In the biotechnology industry, however, ninety percent of the population consists of small and medium size firms and scale economies rarely exist (Audretsch & Stephan, 1996). Accordingly, the number of existing firms in a cluster can properly capture the growth of a cluster.

We consider polynomial regression models to fit the growth patterns of clustered firms. Since hypothesis 1 suggests the growth of geographic clusters evolve in a non-monotonic fashion over time, polynomial regression analysis is applied to test differences in cluster growth across time. The pth order polynomial regression model can be expressed as

$$N_t = a + \sum_{i=1}^{p} t^i + e_t \tag{1}$$

where Nt: total number of biotechnology firms in clusters at year t, a: intercept, and et: error terms.

The first to the third order polynomial regression models, in particular, are considered to see the changes in the growth rates of clustered firms in the industry. Model and coefficient significances and improvement of the goodness of fit can suggest the better fitting of our data trends, testing hypothesis 1. Possible concerns about the multicolinearity among polynomial components can be corrected through the Huber/White/Sandwich correction process in STATA (statistical software that we use).

Univariate/multivariate analysis. The six hypotheses proposed by Pouder & St. John (1996) are related to see the differences in cost economies of resource procurement and legitimacy, the rate of growth in the number of firms, and the innovations between clusters and non-clusters during different phases of the evolution of clusters. We use both univariate and multivariate analysis. For univariate analysis, we conduct mean difference t-tests between clustered firms and non-clustered firms. We assume that variances within two groups are equal, and hence we use a pooled variance to derive t-statistics. As such, the mean difference test statistic that follow t-distribution with (n_c+n_n-2) degrees of freedom can be expressed by (2)

$$t = \frac{\overline{Xc} - \overline{Xn}}{\sqrt{(\frac{1}{n_c} + \frac{1}{n_n})s^2}} \approx t_{n_c + n_n - 2},$$

where $\overline{Xi}$: i's mean of X, n_i : i's number of observations, $s^2 = \frac{(n_c - 1)s_c{}^2 + (n_n - 1)s_n{}^2}{n_1 + n_2 - 2}$, and i = c (for clustered firms) or n (for non - clustered firms).

To complement univariate analysis, we consider multivariate analysis. Event history analysis is undertaken to see how the likelihood of events changes between clusters and non-clusters during different phases, even after controlling industry factors. As such, the event history analysis model can be express as

$$\lambda(t) = \exp(aX(t))$$
$$where \lambda(t) = \lim_{\Delta \to 0} \frac{q(t, t+\Delta)}{\Delta}. \qquad (3)$$

The exponential distribution assumption is suitable for modeling data with a constant hazard and when there is no a priori expectation as to the nature of distribution. Parameters are estimated by the maximum likelihood method using STATA. As an exception, for the average growth rates of clustered and non-clustered firms (hypothesis 3 and 6), we use ordinary least square regression models for multivariate analysis.

Cost economies and legitimacy. To see the effects of agglomeration on cost structure and relations in the relevant community (hypothesis 2 and 5), we consider differences in cost economies and legitimacy between clustered and non-clustered firms. According to the theory by Pouder & St. John (1996), cost economies in resource procurement and obtaining legitimacy are critical factors to the shifts in the economies and diseconomies of agglomeration over time. Since legitimacy is closely tied to efficiency in resource procurement (Hannan & Carroll, 1995: 25), we consider measures for cost economies in resource procurement also reflect the legitimacy concerns. As measures for these concepts, we use the count of public offerings by a sample firm, the cumulative amount of public offerings raised by a sample firm, the count of private offerings by a sample firms, and the count of a sample firm's research alliances. The data on these variables are gathered from the announcements in the Bioscan directory and the Action database by NCBC.

Average growth rates of clusters and non-clusters. To test hypothesis 3 and 6 on the different growth patterns of clusters and non-clusters, we consider the average growth rates of clusters and non-clusters. Since we have multiple regions within clusters and non-clusters, we use the average growth rates of two categories. To measure the average growth rates of clusters and non-clusters, we simply regress the total number of sample firms in each category on years of our study period. And we use the estimated betas from the simple regression as the measures for the average growth of clusters and non-clusters. The empirical model can be expressed by

$$N_{it} = a + bt + e_{it} \quad (4)$$

where N_{it}: total number of firms in I (clusters or non-clusters) at year t (year from 1973 to 1997).

Innovation. For testing hypothesis 4 and 7, we measure innovation performance as the count of patents held by a sample firm. The count of patents is frequently used as a measure for innovation performance (e.g., Audretsch & Feldman, 1996; DeCarolis & Deeds, 1999). One possible concern about this measure is related to the fact that many industries do not depend on patents to protect the profits from innovations (e.g., Levin, Klevorick, Nelson & Winter, 1987). As reviewed in the brief industry history, however, the industry shows important roles of patents in protecting the economic values of innovations (Ryan, et al., 1995). Accordingly, we expect relatively little biases from this measure.

Industry variables. To control for the industry variations in the multivariate analysis, we consider measures for industry-level activities prior to the current period and stock market variations. We use total count of public and private offerings by total firms in the sample within the three months prior to the current month. These measures control for "hot" and "cold" financing windows. We also use quarterly total number of research alliances and patents held by total sample firms in three month prior to the current month. Biotechnology stock index (monthly average) is also considered.

Results

Growth patterns of clustered firms. Polynomial regressions fit the growth patterns of clustered firms over time. As exhibited in Table 2, among others, the third order polynomial model fits best the growth of clustered firms over time by its R-square of almost 99.5 percent. As seen in its F-test statistic ($F(1, 21)=125.06$) for the improvement of model significance, the time-cube term is highly significant, implying that the growth rate of clustered firms is declining after its increase up to a certain point (the end of the origination phase). Accordingly, this supports our hypothesis 1. The second order polynomial model shows marginally different from the first order polynomial model with its slightly significant model improvement by the term of time-square ($F(1, 22)=2.96$). The forth order polynomial model only slightly improves its model significance ($F(1, 20)=5.44$), compared to the third order polynomial model (not reported in the table).

[Insert Table 2 about here]

Univariate analysis. Using mean difference test statistics that follow t distribution, we test across-phase changes in difference between clustered and non-clustered firms' cost economies and legitimacy effects (hypothesis 2 and 5), average growth rates (hypothesis 3 and 6), and innovation performance (hypothesis 4 and 7). Table 3 contains the detailed results of our univariate analysis. First, in terms of cost economies and legitimacy, all proxies (average count of public offerings, average count of private offerings, average cumulative amount of financing, and average count of research alliances) exhibit the decrease in the differences between clustered and non-clustered firms over time as p-values of t-statistics decrease across phases. That is, as the phase moves from the origination to the convergence, clustered firms' benefits associated with cost economies and legitimacy decrease, compared to those of non-clustered firms. Significant differences between clustered and non-clustered firms in their private and public funding activities and their alliances still exist in the industry throughout our study period. However, as the hypothesis 2 and 5 predict, it is evidenced that the benefits from geographic clustering in clustered firms' funding and alliance activities decrease over time.

Regarding the difference in the average growth between clustered and non-clustered firms, the mean difference tests suggest that there exists no significant difference in the average growth between clustered and non-clustered firms in the convergence phase, while clustered firms outgrow non-clustered firms during the origination phase. T-statistic for the origination phase is 75.85 significant (away from the 99 percent critical value of 2.75 in the t-distribution with more than 30 degrees of freedom), contrasted with t-statistic of –2.09 insignificant (given the 99 percent critical value of –4.6 in the t-distribution with 4 degrees of freedom) during the convergence phase. This result evidences hypothesis 3 and 6 proposing that the growth of clustered firms will be bounded and converges into industry growth average over time.

Innovation performance also exhibits decrease over phases, supporting hypothesis 4 and 7. As shown in table 3, in the origination phase the average count of patents is significantly different between clustered and non-clustered firms with t-statistic of 10.37 passing the 99 percent critical t-value 2.75 with more than 30 degrees of freedom, while the t-statistic of the mean difference reduces into 5.15 almost by a half given the same critical value during the convergence phase. This significant reduction of p-values between the origination and the convergence phase suggests that innovation performance of clustered firms has decreased over time, implying the magnitude of agglomeration economies dwindle over time.

[Insert Table 3 about here]

Multivariate analysis. In most cases, our multivariate analysis also supports the results from the univariate analysis, even after controlling for possible industry effects. As exhibited in table 4, hazard rate analyses indicate that clustered firms' activities related to cost economies and legitimacy, and their innovation performance diminish over time after controlling for industry variations (average growth of clustered and non-clustered firms are not included in multivariate analysis).

Models for public and private equity offerings suggest that clustered firms benefit more in the origination phase, compared to themselves in the convergence phase and non-clustered firms over time. This confirms that the benefits of access to financial resources within geographic clustering decrease over time. Innovation performance model also exhibits consistent results with univariate analysis. The dominance of innovation performance of clustered firms in the origination phase, compared to that of clustered firms in the convergence phase or non-clustered firms in any phases, signifies that outperforming innovation performance of clustered firms reduces over time. However, models for research alliances and total amount of financing show increase in the relevant activities of clustered firms over time, after controlling for industry effects. These results are inconsistent with the results of univariate analysis. We may need to further elaborate the models with respect to alliances and total financing amount.

Industry control variables shows positive signs, as we expected. That is, as industry levels of all events (public and private equity offerings, research alliances, and innovations) increase, an average biotechnology firm is more likely to experience the events. The negative sign of bio-stock index in the model of private offerings shows the substitutable relation between public and private offerings. The negative sign of bio-stock index in the innovation model may imply that there exist lags between financial funding and innovation outcomes (patenting).

[Insert Table 4 about here]

Concluding remarks

We have empirically tested Pouder & St. John's propositions (1996) on the evolutions of geographic clusters, in the context of the U.S. biotechnology industry. We find that clustered biotech firms exhibit higher cost economies and legitimacy in obtaining resources (e.g., financial funds by private and public equity offerings, and strategic alliances) than non-clustered competitors during the origination period of cluster development. While, after the early period, such clustered firms' advantages in resource access are statistically significantly declining. Compared to that of non-clustered firms, innovation performance of clustered firms shows similar declining over time. In addition, the difference of growth in number of firms between geographically clustered and non-clustered regions decreases after the origination period. These findings suggest the decline of positive net benefits from geographic clustering over time. In particular, the evolutionary path of clustered firms in our sample appears to follow the third-order polynomial function of time.

Given the inconsistent empirical results in the literature on whether there are positive net benefits to geographically clustered firms, this study sheds light on the importance of time dimension in geographic clustering to understand the net benefits of geographic clustering. Since economies and diseconomies of agglomeration change over time, the net benefits of geographic clustering can be time-variant, possibly leading to different empirical results if the evolution of geographic clustering is not appropriately considered. Accordingly, this study answers a call to empirical studies on dynamics of geographic clustering.

This study is not free from its limitations. Some concerns may be related to right-censoring issues. As we see that the average length of the convergence phase in our sample is three years while that of the origination phase is twenty-two years, there may be concerns about small number of sample years of the convergence phase. More year observations in updated data sets should be considered in the future study. For more generality, we can also consider other industries in which cluster declines are already observed (e.g., decline of Route 128 Boston areas in the minicomputer industry, as suggested by Pouder & St. John (1996)).

REFERENCES

Almeida, P., & Kogut, B. 1999. Localization of knowledge and the mobility of engineers in regional networks. **Management Science**, 45: 905-917.

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

Audretsch, D. B., & Stephan, P. E. 1996. Company-scientist locational links: The case of biotechnology. **American Economic Review**, 86; 641-652.

Bania, N., Calkins, L. N. & Dalenberg, D. R. 1992. The effects of regional science and technology policy on the geographic distribution of industrial R&D laboratories. **Journalof Regional Science**, 32: 209-228.

Barney, J. B. 1991. Firm resources and sustained competitive advantage. **Journal of Management**, 17: 99-120.

Baum, J. B., & Haveman, H. 1997. Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990. **Administrative Science Quarterly**, 42: 304-338.

Bruderl, J. & Schussler, R. 1990. Organizational mortality: The liability of newness and adolescence. **Administrative Science Quarterly**, 35: 530 – 547.

Cameron, A. & Trivedi, P. 1990. Regression based tests for overdispersion in the Poisson model. **Journal of Econometrics**, 46: 347-364.

DeCarolis, D. M., & Deeds, D. L. 1999. The impact of stocks and flows of organizational knowledge on firm performance: An empirical investigation of the biotechnology industry. **Strategic Management Journal**, 20: 953-968.

Eisenhardt, K. M. & Schoonhoven, C. B. 1990. Organizational growth: Linking founding team, strategy, environment and growth among U.S. semiconductor ventures, 1978-1988. **Administrative Science Quarterly**, 35: 530-547.

Ellison, G., & Glaeser, E. L. 1999. Geographic concentration of industry: Does natural advantage explain agglomeration? **American Economic Association Papers and Proceedings**, 89: 311-316.

Gersbach, H., & Schmutzler, A. 1999. External spillovers, internal spillovers and the geography of production and innovation. **Regional Science & Urban Economics**, 29: 679-696.

Greene, W. H. 1993. **Econometric analysis** (2nd ed.). Macmillan: New York.

Hill, J. & Naroff, J. L. 1984. The effect of location on the performance of high technology firms. **Financial Management**, 13: 27-36.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. **Quarterly Journal of Economics**, 108: 577-598.

Krugman, P. 1994. Location and competition: Notes on economic geography. In Rumelt, R. P., Schendel, D. E., & Teece, T. J. (Ed.) **Fundamental issues in strategy: A research agenda**: 463-493. Harvard Business School Press.

Lomi, A. 1995. The population ecology of organizational founding: Location dependence and unobserved Heterogeneity. **Administrative Science Quarterly**, 40: 111-144.

Marshall, A. 1920. **Principles of Economics** (8th ed.). Macmillan: London.

Porter, M. E. 1990. **The Competitive Advantage of Nations**. Free press: New York.

Porter, M. E. 1998. Clusters and the new economies of competition. **Harvard Business Review**, 76: 77-90.

Pouder, R., & St. John, C. H. 1996. Hot spots and blind spots: Geographical clusters of firms and innovation. **Academy of Management Review**, 21: 1192-1225.

Prevezer, M. 1997. The dynamics of industrial clustering in biotechnology. **Small Business Economics**, 9: 255-271.

Saxenian, A. 1994. **Culture and competition in Silicon Valley and Rout 128**. Cambridge, MA: Harvard University Press.

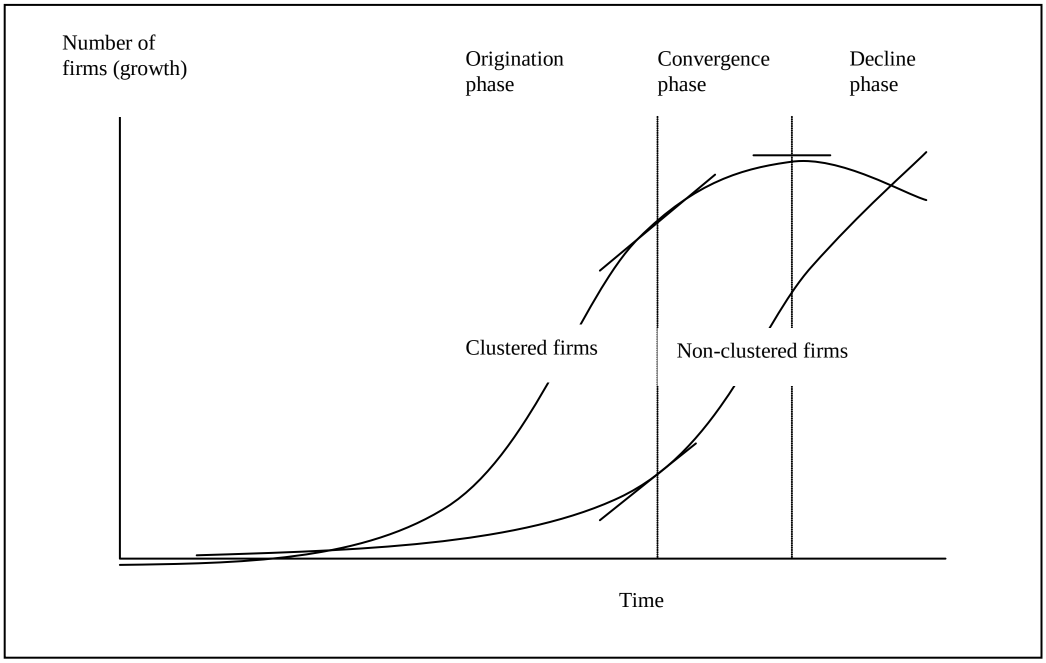
Scott, A. J. 1989. **New industrial spaces: Flexible production organization and regional development in North America and Western Europe**. London: Pion, Ltd.

Staber, U. 1998. Organizational survival in small-firm clusters. **Academy of Management Proceedings**: ENT: A1-A7.

Zucker, L. G., Darby, M. R., & Brewer, M. B. 1999. Intelletual human capital and the birth of U.S. biotechnology enterprises. **American Economic Review**, 88: 290-306.

Figure 1

Evolutionary path of clustered and non-clustered firms*



* This figure is modified from the one by Pouder & St. John (1996).

Table 1

Features and evolution of geographic clusters in the biotechnology industry, 1973-1997*

	Total Entries	Total Exits	Total Density	Industry cumulative density (%)	Year of the first entry	Estimated end of the origination phase**	Convergence phase	Length of origination phase (years)	Length of convergence phase in the study period (years)
San Francisco (Cluster 1)	136	39	97	97 (15.7)	1973	1994	1994-1997	22	3
Boston MA (Cluster 2)	100	26	74	171 (27.7)	1973	1994	1994-1997	22	3
San Diego (Cluster 3)	74	14	60	231 (37.4)	1973	1995	1995-1997	23	2
NY-Tristate (Cluster 4)	73	19	54	285 (46.1)	1974	1993	1993-1997	20	5
D.C. area (Cluster5)	54	10	44	329 (53.2)	1973	1994	1994-1997	22	3
Philadelphia (Cluster 6)	53	12	41	370 (59.9)	1975	1995	1995-1997	21	4
L.A. area (Cluster 7)	37	16	21	391 (63.3)	1973	1993	1993-1997	21	4
N. Carolina (Cluster 8)	37	6	31	422 (68.3)	1973	1995	1995-1997	23	2
Texas (Cluster 9)	30	6	24	446 (72.2)	1976	1994	1994-1997	19	6
Seattle (Cluster 10)	29	6	23	469 (75.9)	1978	1996	1996-1997	19	6
Wisconsin (Cluster 11)	15	1	14	483 (78.2)	1977	-	-	-	-
Colorado (Cluster 12)	14	6	8	491 (79.4)	1980	-	-	-	-
Minnesota (Cluster 13)	14	5	9	500 (80.9)	1975	-	-	-	-
Michigan (Cluster 14)	14	6	8	508 (82.2)	1978	-	-	-	-
Florida (Cluster 15)	13	2	11	519 (84.0)	1979	-	-	-	-
Oregon (Cluster 16)	11	1	10	529 (85.6)	1974	-	-	-	-
Ohio (Cluster 17)	11	3	8	537 (86.9)	1973	-	-	-	-
Georgia (Cluster 18)	10	1	9	546 (88.3)	1981	-	-	-	-
Utah (Cluster 19)	10	2	8	554 (89.6)	1973	-	-	-	-
Other non-clustered regions	90	26	64	618 (100)	1973	-	-	-	-
Industry total/average	825	207	618	618	1973	1994	1994-1997	22	3

*Since we treat smaller clusters than cluster 10 in Seattle area as non-clusters, phases of clusters in non-cluster areas cannot be considered by definition. By definition, the convergence period ends when the difference between the density of a cluster and the average density of non-clusters becomes zero. During our study period, the industry exhibits no such points, leading to only origination and convergence phases in the industry.

** We estimate the end of the origination phase of each cluster by the largest difference between density of a cluster and the average density of non-cluster regions (i.e., cluster 11 to 19 and other non-clustered regions).

Table 2

Polynomial regression models on the growth rates of clustered firms, 1973-1997

	1st order polynomial model	2nd order polynomial model	3rd order polynomial model
Time	25.90 (1.38)	34.37 (5.74)	-15.72 (4.25)
Time-square		-0.36 (0.21)	4.36 (0.42)
Time-cube			-0.12 (0.01)
Constant	-58.29 (16.87)	-100.92 (34.91)	18.11 (8.35)
R-square	0.9495	0.9579	0.9947
F statistics		F(1, 22)=2.96	F(1, 21) = 125.06
N	25	25	25

Table 3

Univariate analysis: mean differences between clustered and non-clustered firms over phases*

	Origination			**Convergence**			**Overall**		
	Clustered firms	Non-clustered firms	t statistic	Clustered firms	Non-clustered firms	t statistic	Clustered firms	Non-clustered firms	t statistic
Cost economies and legitimacy									
Yearly firm avg. # of public offerings	1.03 (1.29)	0.61 (0.88)	12.17	1.30 (1.36)	0.84 (1.07)	7.18	1.11 (1.32)	0.67 (0.94)	14.57
Yearly firm avg. # of private offerings	2.13 (2.65)	0.97 (1.61)	16.27	3.20 (3.11)	1.94 (2.60)	8.64	2.46 (2.85)	1.22 (1.95)	18.78
Yearly firm avg. amount of financing	48.66 (91.44)	10.99 (34.61)	15.76	61.60 (89.82)	24.94 (36.70)	9.09	52.70 (91.13)	14.51 (35.66)	18.55
Yearly firm avg. # of research alliances	5.17 (7.40)	2.37 (3.57)	14.34	6.36 (8.09)	2.88 (3.73)	9.53	5.54 (7.64)	2.50 (3.62)	17.50
Innovations									
Yearly firm avg. # of patents	6.03 (11.94)	2.75 (6.40)	10.37	6.17 (12.62)	3.21 (7.92)	5.15	6.08 (12.16)	2.86 (6.82)	11.52
Total firm years (N)	6959	1510		3164	509		10123	2019	
Average growth rate									
Estimated beta of a regression on time	27.70 (1.05)	9.03 (0.48)	75.85	-9.00 (5.20)	-2.5 (1.44)	-2.09	24.9 (1.20)	7.71 (0.55)	65.11
Total years (N)	22	22		3	3		25	25	

* In the parentheses, standard deviations are provided. Critical values of t-statistics with +100, 40 and 4 degrees of freedom at 0.01 significance level are 2.58, 2.71 and 4.60, respectively.

Table 4

Multivariate analysis: Results from hazard rate and OLS regression models

	Public offerings	**Private offerings**	**Research alliances**	**Innovations (patents)**	**Amount of financing (OLS)**
Clustered firms in the origination phase (=1)	1.07*** (0.16)	1.05*** (0.17)	0.68*** (0.15)	0.82*** (0.24)	34.80*** (8.31)
Clustered firms in the convergence phase (=1)	0.23 (0.17)	0.69*** (0.16)	0.80*** (0.15)	-0.17 (0.26)	44.42*** (7.47)
Industry total in quarterly public offerings	0.02*** (0.007)				
Industry total in quarterly private offerings		0.02*** (0.003)			
Industry total in quarterly research alliances			-0.01*** (0.001)		
Industry total in quarterly patents				0.02*** (0.001)	
Industry (Bio) stock index	0.14* (0.06)	-0.3*** (0.93)	-0.07 (0.04)	-0.20*** (0.04)	4.35 (2.42)
Constant	-9.89*** (0.26)	-8.04*** (0.22)	-6.88*** (0.18)	-7.93*** (0.30)	-0.22 (8.41)
Likelihood/R-square ^	-1172.87	-1582.97	-969.67	743.85	0.0345
N	11311	11311	11311	11311	12160

^ Since the model of amount of financing is estimated by ordinary least square estimation, R-square is provide.