Increasing Returns to Economic Activity Concentration

Fernando Buendía
Business Administration Department
University of the Americas, Puebla
Santa Catarina Mártir, Cholula, Puebla, 72820 Mexico
fbuendia@mail.udlap.mx

for the
23rd International Conference of the System Dynamics Society
Boston, Massachusetts, USA
July 17-21, 2005

Abstract

This paper introduces the notion of increasing returns to economic activity agglomeration and develops a formal system-dynamic model where this notion is used to explain the self-organizing nature of the spatial structure of industrial clusters. In this model, both pecuniary and external economies based on knowledge spillovers are considered.

1. Introduction

Presently, economic geography has an important place in economic theory. Its study has revealed that geographic space is essential to understand the divergence in the allocation of economic activity among regions, and eventually the difference in economic growth of the countries where those regions are located. In the growing literature on industrial clusters two different strands can be identified. One of them comprises various case studies of flourishing regions, such as Silicon Valley, the Third Italy, Baden Würtemberg, and Route 128-Boston. These case studies identify the specific reasons for these clusters' success, but they do not provide a general explanation of how industrial clusters evolve. The second strand of literature has gone beyond the simple description of factors causing successful regions and developed some general concepts to explain the origin of *agglomeration economies*, which has traditionally thought of as the main driving force behind the self-organizing patters of location of firms in different geographical areas. Within this part of the literature there are two recurrently quoted theories of why regions concentrate economic activities. One of these theories is the highly influential interpretation of regional blocks developed by Porter (1990). Rooted in the strategic management tradition, this approach stresses the positive effects of "agglomeration economies" —that come from the interacting elements of the competitive diamond: local demand, infrastructure, government, communication technologies, availability of inputs, and access to output markets— on the firm's capabilities to compete in international markets.

Although Porter's Model has an important place in the literature and is the foundation of important empirical studies —see, for instance, the Global Competitiveness Report developed by Porter et al. (2002)—, its theoretical relevance has been questioned by Krugman (1994), who argues that Porter's model is hard to handle for an economists. Independently of what Krugman means by such an argument, it is clear that the approach of Porter is more a list of variable than a real theory. Obviously, this list of variables is helpful to characterize the phenomenon such variables are related to, but it is still far from being a formal theoretical explanation. Krugman's (1991a, 1991b, 1996a 1996b) integrated the spatial dimension into main stream economic theory. To explain why a particular industry is concentrated in a particular area, he relies on a model based on *agglomeration economies*. Krugman (1991b), invoking Marshall (1920), argues that these economies arise from three main factors: a pooled labor market, pecuniary externalities enabling the provision of inputs to an industry in a greater variety and at lower cost, and technological spillovers. In Krugman (1991a), he develops a different model to answer a different question. Instead of asking the reason why a *particular* industry is concentrated in a particular area, he asks himself why the location of manufacturing in *general* exhibits properties of self-organizing systems, which can be described by a simple power law: manufacturing ends up concentrated in one or few regions of a country, with the remaining regions playing a marginal role of raw material supplier to the manufacturing center. To address this question, he makes a distinction between pecuniary externalities (as when the actions of one firm affect the demand for the product of another firm whose price exceeds marginal cost) and technological spillover (as when the research and development efforts of a firm spill over into general knowledge pool). For Krugman the pecuniary externalities are more "real" than technological spillovers. Therefore, he believes that the use of pecuniary externalities makes the analysis much more concrete than if external economies arise in some invisible form.

With this distinction in mind Krugman (1991a) asks himself where the manufactures production will take place in the presences of pecuniary externalities. If scale economies are assumed, production of each manufactured good will take place at a limited number of sites. *Ceteris paribus*, the preferred sites will be those with relative nearby demand. But some of the demand for manufactures will come from the manufacturing sector itself. Krugman (1991a) argues that this gives room to what Myrdal (1957) called "circular causation" and what Arthur (1989) has called "positive feedbacks": manufactures production will tend to concentrate where there is a large market, but market will be large where manufactures production is concentrated. This "backward linkage" may be reinforced by a "forward linkage": *ceteris paribus*, it will be more desirable to produce near a concentration of manufacturing production because it will then be less expensive to buy the goods this central place provides. Krugman (1991a) also argues that with mass production and lower transportation costs, the tie of production to the distribution of land will be broken: a region with a relatively large nonrural population will be an attractive place to produce both industrial and agricultural goods because of the large local market and because of the availability of the goods and services produced there. As Krugman himself recognizes, this is not an original idea; in fact, a story along these lines has long been familiar to economic geographers, who emphasize the role of circular processes in the emergence of the U.S. manufacturing belt in the second half of the nineteenth century. What is original about Krugman's model is that he embodied this story in a simple yet rigorous model.

Even if this model has been a very valuable reference to explain and measure the levels of concentration of the economic activity in various regions —Krugman's (1991) calculation of gini-coefficients is the main example of this kind of measures—, he has ignored two important aspects concerning industrial clustering. In the first place, instead of taking into consideration a number of important variables which are crucial to the formation and evolution of industrial clusters, such as urban infrastructure, entrepreneurship, innovation, R&D investment, the growth and increasing efficiency of the firm and so on, he has focused in an extremely reduced number of variables: in one of his models he considers pool of labor, provision of inputs and spillovers as the main variables (see Krugman, 1994), and in the other model he sees demand, low transportation costs and scale economies as the most important variables (see Krugman, 1991a).

Another aspect that Krugman (1991b) has neglected has to do with the mutual causality among variables. Conventionally, agglomeration economies have been thought of as coming exclusively from technological spillovers. The existence of technological spillovers, obviously, entails that some kind of mutual causality is in operation. However, little effort has been done to determine how and what variables affect each other to cause such spillovers. Consequently, technical spillovers produced by a mutual causality have been more assumed that explained.

Krugman (1991a) talks explicitly about mutual causality, however, he hardly specifies how and what variables affect each other to produce pecuniary externalities. As a matter of fact, he relies heavily on low transportation costs and strong scale economies to explain manufacturing activity concentration. It is obvious that these economies are important, as I will discuss later, but they are just part of the whole story. As a matter of fact, economic activity concentration is the result of a more complex array of interrelations among numerous variables and not just from scale economies. Obviously, these models have shed light on the fundamental questions of economic geography, but they lack the explanatory power to deal with the complex interactions among variables that in fact takes place in real-world industrial clusters.

Buendía (2004) has pointed out the weaknesses of Krugman's model and suggests that the application of systems dynamics and urn theory to economic geography theory can help to provide stronger and more general conclusions about the emergence and evolution of industrial clusters, for the simple fact that they are *complex systems*. As in other complex system —Buendía (2004) argues— the emergence and evolution of industrial clusters are not the result of the activity of a fist of actors or question of a simple cause-effect relationship between two or three variables, not matter how important these variables might be.

The purpose of this paper is to develop a formal model of geographical concentration of economic activity based on the notion of *increasing returns to economic activity agglomeration*. The interesting aspect of this model is that the concentration of economic activity in specific locations depends on the mutual causality among many variables. This paper has three more sections. Section two describes in an informal way how increasing returns to economic activity agglomeration are produced. In section three I develop the analytical model. The paper ends with some conclusions and final comments.

2. Sources of Increasing Returns to Economic Activity Agglomeration.

While conventional industrial clusters literature has determined that geographic space matters for industrialization, urban development and economic growth, it has yet to unravel in a more detailed way how agglomerations are formed, where they come from, how they are either sustained and strengthen, or else deteriorate over time. In other words, it has to go beyond the simple list of variables affecting industrial clusters and the static models consisting of two or three variables whose interaction is practically ignored to give place to increasing returns to economic activity agglomeration. The fact that economic activity concentration is produced by increasing returns is not new. However, they usually are associated to increasing returns to scale in production. Krugman (1991a), for instance, argues that with lower transportation costs, a higher manufacturing share, or stronger economies of scale circular causation sets in and manufacturing will concentrate in whichever region gets a head start.

This account of industry localization is extremely relevant, but it relies heavily on scale economies to production. It is clear that scale economies play an important role in economic activity location, but there are many other important variables that may affect the emergence and evolution of industrial clusters. There is empirical evidence that shows that concentration of economic activity in specific locations emerges not only from the fruitfully cooperation and interaction of a large number of economic actors and institutions, but also from the multifaceted arrangement of relations that result from the mutual causality among numerous variables. Thus it seems accurate to thought of industrial clusters as complex systems. Buendía (2004) suggests that, as any other dynamic systems, industrial clusters are subject to both negative and positive feedbacks. Negative feedbacks produce *decreasing returns to economic activity agglomeration* —reductions of benefit due to the increasing number of other firms—, which may occur because locations become congested, land become expensive, or because infrastructure may become scarce and costly. Decreasing returns to economic activity agglomeration are stabilization forces which delay the growth of industrial districts and prevent the eventual emergence of an infinite-size cluster.

The growth of industrial clusters depends to great extent on positive feedbacks, which produce *increasing returns to economic activity agglomeration*. This kind of source of increasing returns are different from the conventional agglomeration economies, for *increasing returns to economic activity agglomeration* the are related to a complex assembly of interaction that result from the mutual causality among several variables, while agglomeration economies are related to knowledge spillovers and scale economies, as in the conventional models analyzed before. Concerning industrial clusters, many interactions and mutual causality among variables may cause increasing returns to economic activity agglomeration, but from the consistent results found in the literature, Buendía (2004) identifies the following as the most important:

Competitive Advantage, economic growth and manufacturing concentration. According to Porter (1990) industrial clusters improve countries' competitive advantage and export position, which in turn propels still further the economic growth of those countries. Then at the highest level of generalization a mutual causality can be established between cluster growth, competitive advantage, and export position improvements.

Industrialization and urbanization. Industrial clustering is strongly correlated with urbanization. Manufactures production will tend to concentrate where urbanization is well developed, but where urbanization will be large where manufactures production is concentrated. Kim and Margo (2003) have shown that, while cities existed in the pre-industrial era, the rapid growth in the number and size of cities coincided with the accumulation in those cities of manufacturing activities, especially capital and knowledge-intense industries. During this period, rural areas urbanized more rapidly than the developed urban areas. Furthermore, urban growth was consistently unrelated to the initial size of cities and younger cities grew faster than older cities. The fact that industrialization is seen to arise in rural areas suggest that industrialization preceded urbanization, but in the urban development literature is also accepted that urbanization encourages industrialization.

Clusters, knowledge and spillovers. Traditionally, spillovers have thought of as important factors that enhance the innovativeness and cost efficiency of firms in clusters. The existence of local spillovers implies that firms in a cluster can profit from innovations and technological improvements developed by other firms in the same cluster. Spillovers occur if an innovation or improvement implemented by a certain enterprise increases the performance of another enterprise without the latter benefiting enterprise having to pay (full) compensation. Larger stocks of knowledge lead also to more innovations, and son on. Feldman (1994a) has provided empirical evidence that what is true for production was even more pronounced for innovative activity.

Clusters, innovations and firms. The relationship between innovations and clusters would not be complete without an analysis of the nature of firms in clusters. The evidence shows that innovations are largely concentrated (Feldman, 1994a). In the United States, for instance, only 11% of the states accounted for 81% of all innovations. Usually, innovations are positively influenced by both industrial R&D expenditure and university research. However, university research is more important for small firms than large firms, whereas corporate R&D stimulates more innovations in large firms more than in small firms. Therefore, clusters are especially important to small firms, which make them likely to support the development of radically new products and technologies. Beaudry and Breschi (2000) advance a similar idea. They analyze whether in a strong industrial cluster really facilitates firms' innovative activities. Whereas location in a cluster densely populated by innovative companies in a firm's own industry positively the likelihood of innovating, quite strong disadvantages arise from the presence of non-innovative firms in a firm's own industry. Therefore, positive knowledge externalities are likely to flow from innovative companies. Clusters having a higher number of innovative companies and a larger stock of knowledge in the past are more likely to a better innovative performance. Furthermore, Caniëls and Romijn (2003) have argued that a close look at the behavior of the individual actors (notably innovative firms) that make up a region is essential for gaining a better understanding of the driving forces of regional dynamism.

An increase in the number of innovations is not the only effect firms can expect from being in a cluster. Clusters will also lead the firm to an increased competitive position in the international market. Since firms in a cluster are supposed to exploit economies of scale and scope, they reduce their cost which increases their profits and strengthens their competitive position. There is also empirical evidence that shows that industrial concentration, firm size, and market share are related to the number of innovations in clusters. Then large clusters generate more innovations. There also seems to be a positive relationship between market share and innovativeness, which entails a feedback between innovative success and market power. Therefore, high levels of industrial concentration lead to more innovations as well.

Clusters, innovations and the firm's profitability. An important question in the economic geography literature is whether cluster lead to higher profits. This relation is not easy to find directly, but it may also be studied indirectly by calculating the effects of innovations on profits. If a firm in a cluster innovates more than the same firm outside the cluster, this implies that it also spends more on R&D. Since the studies show that R&D is profitable for a firm as well as for the society as a whole, it may be concluded that participating in a cluster increases the profits of a firm and generates social benefits that result in higher economic growth.

Clusters, knowledge institutions and skilled workers. Knowledge spillovers and the presence of skilled labor are important cluster determinants. Surprisingly, none analyses reviewed by Hoen (2001) found that cooperation between firms and knowledge institutions in a cluster matters, even though such cooperation is usually assumed to lead to considerable gains for firms. Actually, Beugelsdijk and Cornet (2001) found no evidence that proximity matters for spillovers in the Netherlands and that the presence of a university of technology is positively related to the innovativeness of neighboring firms. Therefore, research companies and innovation centers are less important as external source of information and innovations to a firm than related firms, suppliers, and buyer.

Given that most cooperation agreements are between firms and not between firms and research or educational organizations, knowledge institution cannot be considered a cluster determinant. This does not mean, however, that institutions such as universities do not play a role in the existence of knowledge spillovers, since they still publish results of their research, provide services and educated work force for firms. According to Hoen (2001), most authors argue that the government should provide for the infrastructure which makes the emergence of clusters possible. Since clusters need skilled labor, a good cluster policy should support a university system that produces a highly educated and skilled workforce. Education, therefore, is still a cluster determinant.

Clusters and infrastructure. To study clusters in an integral way, we have to consider the fact that they are embedded in an urban structure. Relaying on recent insights on the importance of "organizing capacity", Van den Berg et al. (2000) take urban aspects into account to study the growth process of clusters. They develop a framework containing the following interrelated elements: economic and spatial conditions, life quality, and organizing capacity of urban regions. The economic conditions of an urban region comprise the state of the local and international demand, which is fundamental to the functioning of clusters. In general, quality of life is an essential locational factor in the economic development of urban areas. It can include services which increase the quality life of people, such as education, entertainment, and health service. Highly skilled people, on whom urban development is strongly dependent, attach much value to the quality of the life of the geographical area where they work. The spatial conditions for urban development relates to accessibility, in the broadest sense of the word. Accessibility is a necessary condition for the interaction of firms in a cluster. Bad transport in urban areas, for example, may seriously hamper the interaction in a cluster, particularly if clusters elements are dispersed. External accessibility to other cities and regions through e-mail, telephone, airports, railroads, and harbors is a fundamental means to link local firms with national and international firms and organizations of all kinds. The final element for the development of cluster has to do with the organizing capacity of cluster participants, which includes the presence of a vision and strategy regarding the cluster, a good level of public-private cooperation, political and societal support for cluster development, and leadership.

Figure 2 summarizes our analysis and provides graphical details of the main interdependences that play a role in the emergence and evolution of industrial clusters. Others interdependences can be found among the variables and the behavior of clusters actors, nevertheless the main interactions described in this section seem to be sufficient to show that industrial clusters are in fact dynamic systems and, consequently, that increasing returns to economic activity agglomeration come from the mutual causality among a large number of variables and not just from two o three disconnected parameters. If it is accepted that the geographic concentration of economic activity depends on the abundance of correlations and feedbacks, then we now can count on a safe reference to revise the conventional models. Another eventual benefit of this model it is that it will provide a foundation to richer empirical studies. Traditionally, the spatial distribution of economic activity has been measured using data bases related to innovations and employment. Using this model as a reference, empirical studies can be extended to other parameter, such as quality and availability of urban services and number and importance of inputs providers. The next step is to develop a formal model to give a rigorous formulation to the relationships just described.

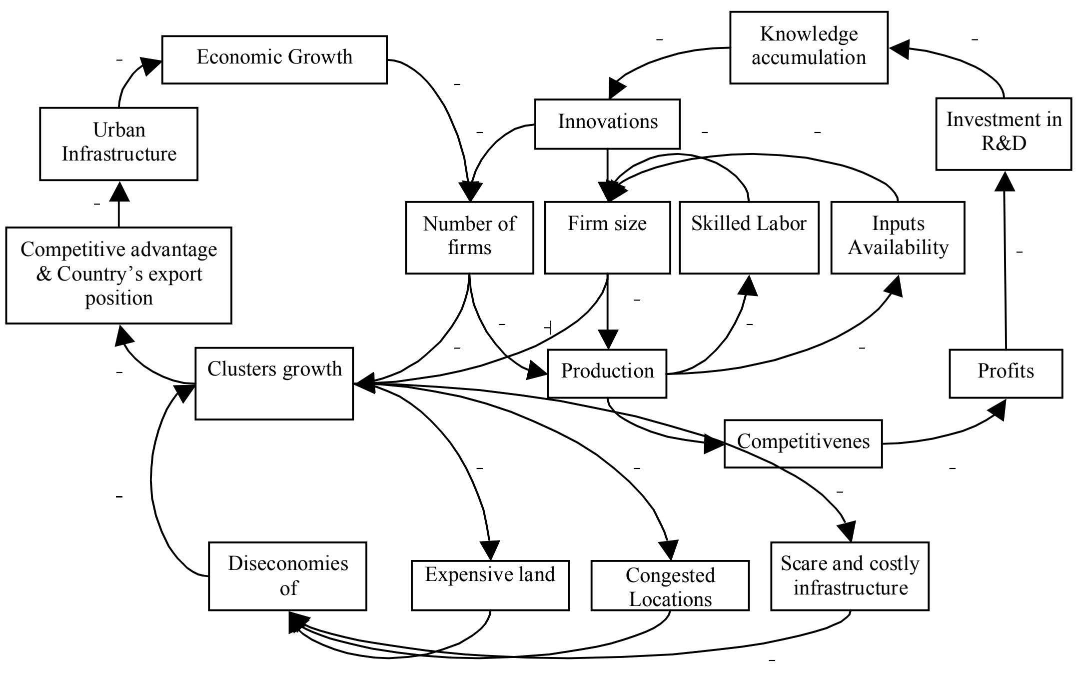

Figure 1. The evolution of industrial clusters

3. The Model

In this model, there is *increasing returns to economic activity agglomeration* as described in the previous section. This implies that a firm's benefits to be in a location together with other firms increase with the number of firms in that location. This is reinforced by the gains in useful infrastructure that is attracted by the larger number of firms. Larger industrial clusters produce larger demand for the products of the firms in the clusters. Larger clusters increase the availability of skilled workforce and inputs, which in turn makes the cluster grow still further. The growing cluster brings about increasing benefits for the firm: it becomes more efficient and competitive, which produces increasing returns to the firms. The firm can invest this returns in R&D with which increases its probability of come about with a new technology. These innovations make the firm either to grow further or produce new firms. These new and growing firms make the size of the cluster still larger, which attract new infrastructure and the cycle feed itself up to the point where just one or few locations stand out with respect to the other locations. Diseconomies of agglomeration —net benefits that decrease with increasing number of firm within the cluster— may also take place. The location may become congested; land may become expensive; infrastructure may become scare and costly. We consider that this process produces increasing returns to economic activity agglomeration, which causes the economic growth of certain geographical areas and the stagnation of others. Figure 2 show the relationships among variables, which are the base of the formal model.

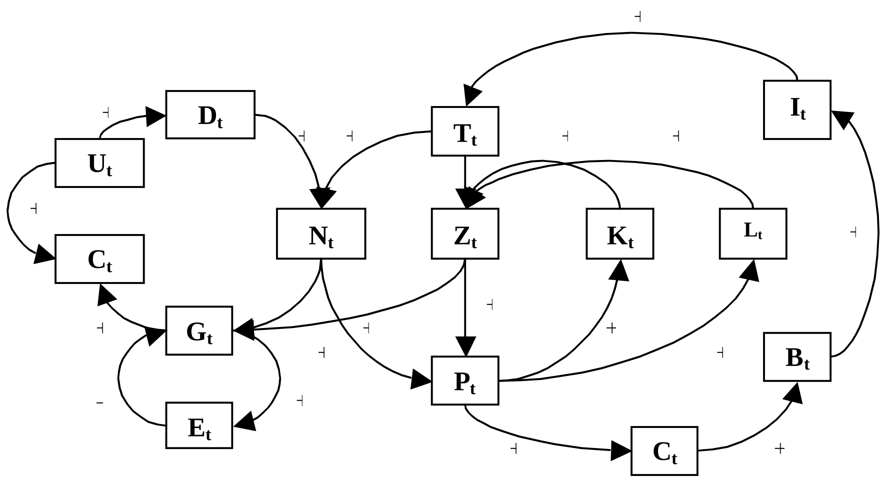

Figure 2. A graphical representation of the model

We can now proceed to build a model where countries grow as their industrial clusters grow. Each country's economic development, D_t^i, depends positively on its competitive advantage (or its export position), C_t^i, and the infrastructure available, U_t^i. Economic development is affected negatively by diseconomies to scale, E_t^i. Therefore the grow, D_t^i, of country *i* at time *t* can be condensed into these components:

$$D_j^i = f\left[C_t^i\left(G_t^i\right), U_t^i(n), E_t^i(n)\right] \qquad (1),$$

where *n* is the number of firms. This function can be rise to the following first-order conditions:

$$\frac{\partial D_t^i}{\partial C_t^i} \succ 0, \frac{\partial D_t^i}{\partial U_t^i} \succ 0, \text{ and } \frac{\partial D_t^i}{\partial E_t^i} \prec 0.$$

The competitiveness, C_t^i, of each country is function of the production level of the clusters in that country, P_t which is the sum of the level of production of each firm. Therefore, the growth of the cluster can be expressed as

$$G_t^i = f\left(P_t\right) \qquad (2).$$

P_t is the total production of all firms over time in the cluster *i*. P_t, therefore, is the sum of each firm's production, *p*, over time:

$$P_t = \sum_{i=1}^{n} p_t^i \qquad (3).$$

The number of firms, n depends on innovations, T, the availability of skill workers, k, and the supply of inputs, L.

$$n_t = f(T, K, L) \tag{4}$$

K and L, however, for the discussion of the previous sections, depend on the number of firms, so to indicate the mutual causality between T, K, L and n, we have to add to our model the following equations:

$$K_t = f(n_t) \tag{5}$$

$$L_t = f(n_t) \tag{6}$$

As we said in our discussion on the system-dynamics character of industrial clusters, new technologies and innovations depend on the total investment of the firms in the cluster, which depends positively on their benefits. We assume that firms in the clusters are subject to increasing returns to production, so benefits increase with production. Therefore, the number of innovation in the cluster is given by:

$$T_t = f\left\{\sum_{i=1}^{n} I_t^i\left[B_t^i\left(P_t^i\right)\right]\right\} \tag{7}$$

As we assume that firms are subject to increasing returns, we can assume that the number of innovations each firm comes up is subject to the following expression:

$$T_t^i = f\left\{I_t^i\left[B_t^i\left(P_t^i\right)\right]\right\} \tag{8}$$

where $\frac{\partial T_t^i}{\partial I_t^i} \succ 0$. Given the above assumptions, we can calculate the probability that a new firm chooses the region j. We can assume that there are $y_1, y_2, \ldots, y_N$ firms in region 1through n. We also assume that a new firm can emerge only when a new technology is introduced. It might be possible that a new technology be introduced by a existing firm, so instead of a new firm, the firm introducing the new technology grows. In any case anew technology increases the production of the firm and the total production of the cluster where it is located. Given that there is a mutual causality between production, profitability and investment, the probability of a region to develop is given by

$$p_j = pro\left\{\left[C_j(G_j), U_j(n), E_j(n)\right] \succ \left[C_i(G_i), U_i(n), E_i(n)\right] \; all \; i \neq j\right\} \tag{9}$$

Arthur (1994) and Arthur et al. (1987) has developed mathematical tools to address this kind of dynamics. He starts with a simple model where firms in an industry decide on locating in one of N possible regions subject to no agglomeration economies. Firms are well informed about the net present value of settling down in each location at the time of their choice. The firms in an industry differ in location needs. In this simple model Arthur condenses all benefits, r_j^i, to firm I for locating in site j, into two components:

$$r_j^i = q_j^i + g(y_j) \tag{10}$$

where q_j^i is the geographical benefit to firm I for setting in location j and $g(y_j)$ is the net agglomeration benefit from having y_j firms already located there at time of choice. Firms are supposed to be drawn of a vector $q = (q_1, q_2, \ldots, q_N)$ which denotes that firm's location tastes for each possible site, from a given distribution of F of potential firms over locational tastes. To compute the probability that an entering firm chooses location j, suppose there are

$y_1, y_2,..., y_N$ firms already in location 1 through *N*, respectively. Then the probability that the next firm "drawn" prefers *j* over all other sites is:

$$p_j = \Pr o\{[q_j + g(y_j)] \succ [q_i + g(y_i)] \ all \ i \neq j\} \quad (11)$$

In our model the

Arthur (1994) modeled the location pattern of firms under increasing returns using a *basic* urn scheme with white and black balls, with each color corresponding to a kind of firm. At its initial state, the urn contains n_w white balls and n_b black balls, and a ball is added at subsequent time instances t = 1, 2, 3, …, N. The probability of this ball being white is given by $f_t(X_t)$, and the probability for a black ball is 1 - $f_t(X_t)$, with the random variable X_t standing for the proportion of white balls in the urn at time *t*. The dynamics followed by the number of white balls w_t depends on a random binary variable $\xi_t(X_t)$, which is independent of time and takes on values from the subset of integer numbers: {1 with *probability* $f_t(X_t)$, 0 with probability 1 - $f_t(X_t)$ }. This dynamics is modeled by

$$w_{t+1} = w_t + \xi_t(X_t), \quad (12)$$

where it is established that the number of white balls at each state remains the same (with *probability*1- $f_t(X_t)$) or it is incremented by one (with *probability* $f_t(X_t)$): $w_{t+1} = w_t$ or $w_{t+1} = w_t + 1$. The *dynamics* that rules the total number of balls γ_t in the urn at time *t* is given by:

$$\gamma_{t+1} = n_w + n_b + t \quad (13)$$

which is incremented by one at each time.

The proportion of white balls X_{t+1} in the urn at time *t+1* is obtained by dividing the number of white balls w_{t+1} by the total number of balls γ_{t+1},

$$X_{t+1} = \frac{w_t + \xi_t(X_t)}{n_w + n_b + t} = \frac{(w_t + \xi_t(X_t))(n_w + n_b + t - 1)}{(n_w + n_b + t)(n_w + n_b + t - 1)} = \frac{(n_w + n_b + t)\, w_t - w_t + (n_w + n_b + t - 1)\, \xi_t(X_t)}{(n_w + n_b + t)(n_w + n_b + t - 1)} \quad (14).$$

In order to have the current value of X_{t+1} expressed in terms of its previous value X_t plus an increment ΔX_t, the following algebraic manipulations have to be performed:

$$X_{t+1} = \frac{(n_w + n_b + t)\, w_t}{(n_w + n_b + t)(n_w + n_b + t - 1)} + \frac{(n_w + n_b + t - 1)\, \xi_t(X_t) - w_t}{(n_w + n_b + t)(n_w + n_b + t - 1)}, \quad (15)$$

$$X_{t+1} = \frac{w_t}{n_w + n_b + t - 1} + \frac{\dfrac{(n_w + n_b + t - 1)\xi_t(X_t)}{n_w + n_b + t - 1} - \dfrac{w_t}{n_w + n_b + t - 1}}{(n_w + n_b + t)}. \quad (16).$$

Since $X_t = \frac{w_t}{n_w + n_b + t - 1}$ and $X_{t+1} = X_t + \frac{\xi_t(X_t) - X_t}{(n_w + n_b + t)}$, then expected value for the increment in X_t is given by the relation:

$$E\left\{\frac{\xi_t(X_t) - X_t}{(n_w + n_b + t)}\right\} = \frac{[1 \cdot f_t(X_t) + 0 \cdot \{1 - f_t(X_t)\}] - X_t}{(n_w + n_b + t)} = \frac{f_t(X_t) - X_t}{(n_w + n_b + t)} \quad (17).$$

The fluctuations at the beginning of the process take a value according to $-1 \le f_t(X_t) - X_t \le 1$, but eventually the fluctuations in ΔX_t tends to zero, and X_t reaches a *steady state* $f_t(X_t) - X_t = 0$. Therefore, it is said that the probability of the event $\left| f_t(X_t) - X_t \right| \approx 0$

converges to 1 as t → ∞ with zero or positive probability; and for an isolated root Φ, the fastness of the convergence of $f_t(X_t) - X_t = 0$ in a neighborhood around Φ, depends on the smoothness of $f_t(X_t)$ at Φ. Other useful way of describing the previous properties of this urn scheme is by defining a function $f(X_t)$ such, that $f_t(X_t) = f(X_t) + \delta_t(X_t)$ and in the *limit* t → ∞, $\delta_t(X_t)$ approaches 0, and $f_t(X_t)$ approaches $f(X_t)$.

This simple urn scheme displays positive feedback and two patterns of evolution reaching a steady state. The behavior of X_t over time describes trajectories with random walks, approaching a limit that can take on any value from the *sub-set* of *real numbers*: [0, 1].

Let us consider a market with two competing products *A* (with $n_A \geq 1$ units) and *B* (with $n_A \geq 1$ units) such that a new consumer enters the market at time instances *t* = 1, 2, …, N The pattern of location of both firms is clearly modeled by the previous urn scheme, where the function $f(X_t)$ is constructed according to the decision rule that the new firms use to make his choice.

As an example, considered the following basic rule: if at least $\frac{r+1}{2}$ firms of the type *A* are located in a region the new firm will choose A, otherwise *B*. The function $f(X_t)$ that represents the probability of the new consumer choosing *A*, depends on the current proportion X_t of product *A* in the market,

$$f(X_t) = \sum_{i=\frac{r+1}{2}}^{r} \frac{r!}{(r-i)!} X_t^{\,i} (1 - X_t)^{r-i} . \tag{18}$$

We are interested in the solution of $f\ (X_t) - X_t = 0$ on $X_t = \in [0,1]$. There are three roots on the sub-set of real numbers [0, 1]: $0, \frac{1}{2}$, and 1; however, there is no possible market structure corresponding to the root $X_t = \frac{1}{2}$; i.e., X_t converges to this root with zero probability as $t \to \infty$. In the other hand, the roots 0 and 1 correspond to possible locational structures, i.e., X_t converges to each of them with positive probability: $X_t \to 1$ which corresponds to the probability for *A* to dominate (being greater than $\frac{1}{2}$) if the initial number of firms n_A of the region *A* is greater than the initial number of firms of region *B*.

Buendía (2003) has developed a generalized urn model which can be suitable to explain combinations of negative and positive feedback, "jumps" and other type of perturbations. Specifically, he extended the classical models called generalized urn schemes beyond the traditional case of multiples independent urns with multiple additions and two colors —these schemes were just described in the previous paragraphs— to the cases where multiple dependent (or independent) urns with several additions (or withdraws), and several colors are considered. For instance, to model de evolution of locational patterns for the case of two regions and multiple firms settling down in a given region at the same instant of time, the previous urn scheme is extended to the case of multiple additions. This scheme considers the addition of $m \geq 0$ white and $n \geq 0$ black balls (or *m* firms in region *A* and *n* firms in region B) at time instance *t*, with *m* and *n* being the values taken by the components of a two dimensional *random* vector $\xi^t(X_t) = [\xi_1^t(X_t), \xi_2^t(X_t)]$, with each component independent of *t*, and both of them not being mutually exclusive, i. e. the occurrence of one does not discard the happening of the other. The *random variable* X_t is the *proportion of white balls* in the urn at time *t*, and the behavior followed by the number of white balls w_1 and the total number of balls γ_1 in the urn over time is given by

$$w_{t+1} = w_t + \xi_1^t(X_t), \tag{19}$$

$$\gamma_{t+1} = \gamma_t + \xi_1^t(X_t) + \xi_2^t(X_t), \tag{20}$$

where the initial conditions of the urn are defined by $w_1 \geq 1$, $b_1 \geq 1$, and $\gamma_1 = w_1 + b_1$. The *proportion of white balls* X_{t+1} in the urn over time is obtained by dividing w_{t+1} by γ_{t+1},

$$X_{t+1} = \frac{w_t + \xi_1^t(X_t)}{\gamma_t + \xi_1^t(X_t) + \xi_2^t(X_t)} = \frac{w_t + \xi_1^t(X_t)}{\gamma_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right]} = \frac{w_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right] - w_t \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t} + \xi_1^t(X_t)}{\gamma_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right]} \tag{21}$$

In order to have the current proportion expressed in terms of its previous value plus the new white balls thrown to the urn Δw_t:

$$X_{t+1} = \frac{w_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right]}{\gamma_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right]} + \frac{\xi_1^t(X_t) - w_t\frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}}{\gamma_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right]} = X_t + \frac{\xi_1^t(X_t) - X_t[\xi_1^t(X_t)+\xi_2^t(X_t)]}{\gamma_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right]}. \tag{22}$$

The probability of the *random vector* $\xi^t(X_t)$ taking the value $\mathbf{l} = \begin{bmatrix} m \\ n \end{bmatrix}$ is given by its *conditional distribution*, $P_t(\xi^t(X_t) = \mathbf{l}) = p_t(\mathbf{l}, X_t)$ and depends on the current value of X_t. The *expected value* for the number of total balls added to the urn at time *t* is

$$E[\xi_1^t(X_t) + \xi_2^t(X_t)] = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}(m+n)\, p_t(\mathbf{l}, X_t) \tag{23}$$

The *expected value* for the proportion of white balls added to the urn at time *t* depends on the proportion of white balls X_t and total number of balls γ_t at time *t*:

$$E\left\{\frac{\xi_1^t(X_t) - X_t[\xi_1^t(X_t)+\xi_2^t(X_t)]}{\gamma_t\left[1 + \frac{\xi_1^t(X_t)+\xi_2^t(X_t)}{\gamma_t}\right]}\right\} = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty}\frac{m - X_t(m+n)}{\gamma_t\left(1 + \frac{m+n}{\gamma_t}\right)}\, p(\mathbf{l}, X_t) \tag{24}$$

Eventually, as $t \to \infty$, the proportion of white balls reaches a *steady state* and its increment $\Delta X_t \to 0$. Thus, $\lim_{t\to\infty} E\{\Delta X_t\} = 0$, implies that $\lim_{t\to\infty} X_t =$ the set of zeros $\in [0,1]$ for which,

$$\lim_{t\to\infty}\sum_{m=0}^{\infty}\sum_{n=0}^{\infty}[m - X_t(m+n)]\, p_t(\mathbf{l}, X_t) = 0. \tag{25}$$

We are interested in the probability of product *A* (product *B*) being adopted by all consumers, i. e. the probability of $X_t \to 1$ ($X_t \to 0$). This probability is modeled by a *steady state* component $p(X_t = 1)$, and a *transient* part $\delta_t(X_t)$:

$$p_t(X_t = 1) = p(X_t = 1) + \delta_t(X_t) \tag{26}$$

The *transient* part vanishes as $t \to \infty$, so that it can be ignored. As an example of how the *steady state* component is modeled, let us consider that a new consumer uses the following rule to decide which product to buy: He asks and $r \geq 3$ adopters which product they use, and if at least αr (α is a real number in (½, 1]) of them use *A*, he will buy *A*; otherwise, he will buy *B*. The probability of the new consumer buying *A* at time *t* is:

$$p(X_t = 1) = \sum_{i=[\alpha r]}^{r} \frac{r!}{(r-i)!} X_t^{\,i} (1 - X_t)^{r-i} \,. \tag{27}$$

Thus, the expression for ΔX_t as t→∞ and $X_t \rightarrow 1$ becomes

$$\lim_{t\to\infty} \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} [m - X_t(m+n)] p\left(\begin{bmatrix} X_t \\ 1 - X_t \end{bmatrix}, X_t\right) = [mp\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}, X_t\right) - X_t(m+n)] p\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}, X_t\right) \tag{28}$$

This is an example of how the tools of the generalized urn model (Buendía, 2003) can be suitable to describe in a formal way dynamic systems such as industrial clusters, no matter the number of variables and the relationships that exit among them. For instance, each urn may be considered a given geographical region. Number of balls can stand for the number of firms. The evolution of the number of balls of different colors describes the size of clusters, the accumulated competitive advantage of the region and the economic growth of the country where these clusters are located.

4. Conclusions and Final Remarks

This paper develops a dynamic model of the emergence and evolution of industrial clusters based on the notion of increasing returns to economic activity agglomeration. From the analysis of this paper we can draw two fundamental lessons. First, the emergence and evolution of industrial clusters is a matter of a large number of variables and interactions among these variables. Second, suitable mathematical machinery to model formally industrial clusters is the generalized urn model.

References

Arthur, W. B. (1994) *Increasing Returns and Path Dependence in the Economy*, Ann Arbor The University of Michigan Press.

Arthur, W. B. (1989) ‘Competing Technologies, Increasing Returns and Lock-in by Historical Events’, *The Economic Journal*, vol. 99, March, pp. 116-131.

Arthur, W. B., Y. M. Ermoliev, and Y. M. Kaniovski (1987) ‘Non-linear Urn Processes: Asymptotic Behavior and Applications’, International Institute for Applied Systems Analysis, Laxenburg, Austria, working paper WP-87-85, 33 p.

Audretsch, D. B and M. P. Feldman (2003) ‘Knowledge Spillover and the Geography Innovation’, to be published in the *Handbook of Urban and Regional Economics*, Volume 4.

Audretsch, D. B and M. P. Feldman (1996) ‘R&D spillovers and the geographic of innovation and production”, *The American Economic Review*, Vol. 86 (3), pp. 630-640.

Beaudry, C. and S. Beschi (2000) ‘Does Clustering really help firms’ innovative activity’, Centro Studi sui Processi di Internazionalizzazione, working paper 111, July.

Beugelsdijk, S. and M Cornet (2001) ‘How far do they reach? The localization of industrial and academic knowledge spillovers in the Netherlands’, working paper No. 2001-47, CentER, August.

Breschi, S. and F. Lissoni (2001a) ‘Localised Knowledge spillovers versus innovative milieux: knowledge “tacitnesss” reconsidered’, in: *Papers in Regional Science*, 80, pp. 255-273.

Breschi, S. and F. Lissoni (2001b) ‘Knowledge spillovers and local innovation systems: a critical survey’, *Industrial and Corporate Change*, Vol. 10 (4), pp. 975-1005.

Buendía, F. (2004) “Towards a Systems Dynamics-based Theory of Industrial Cluster” forthcoming in Karlsson Ch., B. Johansson and R. R. Stough, *Industrial Clusters and Inter-Firm Networks*, Edward Elgar Publisher, London, England.

Buendía, F. (2000) *The Economics of Increasing Returns, the Size of the Firm, Industrial Concentration, and Technological Competition,* Ph. D. dissertation, École des Hautes Études Commerciales, Montreal, Quebec, Canada, November.

Buendía, F. (2003). 'Generalized Urn Theory', working paper, Center for Socioeconomic and Political Studies, University of the Americas, Puebla, Mexico.

Caniëls M. C. J. and H. A. Romijn (2003) 'What drives innovativeness in industrial clusters? Transcending the debate', Working Paper 03.04, Eindhoven Centre for Innovation Studies, Department of Technology Management, Technische Universiteit Eindhoven, The Netherlands, February.

Ciccone, A. and R. E. Hall (1996). "Productivity and the Density of Economic Activity." *American Economic Review*, Vol. 86, No. 1, March, pp. 54-70.

Feldman, M. P. (1994a) *The Geography of Innovation*. Boston: Kluwer Academy Press.

Feldman, M. P. (1994b) "Knowledge Complementarity and Innovation", *Small Business Economics*, October, 6(5), pp. 363-72.

Harris, T. F and Y. M. Ioannides (2000) 'Productivity and Metropolitan Density', unpublished paper, Department of Economics, Tufts University.

Hoen, A. (2001) *Clusters: Determinants and Effects*, Memorandum, CPB Netherlands Bureau for economic Policy Analysis, August.

Jaffe, A. M. Trajtenberg and R. Henderson (1993) 'Geography localization of knowledge spillovers as evidence by patent citation', *The Quarterly Journal of Economics*, Vol. 108, No. 3 August, pp. 577-98.

Kim, S. and R A. Margo (2003) 'Historical perspective on U. S. Economic Geography', Working Paper 9594, National Bureau of Economic Research, March.

Krugman, P. (1991a). "Increasing Returns and Economic Geography." *Journal of Political Economy*, June, 99(3), pp. 483-99.

Krugman, P. (1991b). *Geography and Trade*. Cambridge, Massachusetts: MIT Press.

Krugman, P. (1996a). *The Self-organizing Economy*. Blackwell Publisher.

Krugman, P. (1996b). *Development, Geography and Economic Theory*. Cambridge,

Massachusetts: MIT Press.

Malmberg, A. and P. Maskel (2002) 'The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering', *Environment and Planning*, Vol. 34, pp 429- 49.

Marshall, Alfred, 1920, *Principles of Economics*, 8th ed. London: Macmillan.

Maskell, P. (2001) 'Towards a knowledge-based theory of the geographical cluster', in *Industrial and Corporate Change*, 10 (4), pp. 921-44.

Myrdal, G. (1957) *Economic Theory and Under-developed Regions*, London: Duckworth.

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

Van den Berg, L., Braun, E. and Van Winden, W. (2001) *Growth Clusters in European Metropolitan Cities*, Ashgate.