The Evolution of Localised Industrial Clusters: Identifying the Processes of Self-organisation

Thomas Brenner
Max-Planck-Institute for Research into Economic Systems
Evolutionary Economics Unit
Kahlaische Str. 10
07745 Jena, Germany

ABSTRACT This paper intends to structure the discussion on localised industrial clusters through a classification of processes that may cause their evolution. To this end, the local dynamics and interactions are analysed and those mechanisms identified that cause a self-augmenting process which might lead to the evolution of localised industrial clusters. The usual concepts and several case studies are discussed in the light of this classification.

KEYWORDS: clusters, industrial districts, innovative milieux, self-organisation, evolution.

JEL classification: D21, L60, O18, R30

I want to thank Dirk Fornahl, Ulrich Witt and the participants of the DRUID summer conference 2000, of the seminar at the University of Jena, and of the workshop on “Economic Development (1950-2000): The Italian Case” for stimulating discussions and helpful comments as well as the Germany Federal Ministry of Education and Research for financial support.
1. Introduction

In the last 20 years, regional aspects of economic activity have been put back on the economic agenda. Under the headings of ‘industrial districts’, ‘industrial clusters’, ‘innovative milieu’ and ‘regional innovative systems’, it has been studied why certain regions are successful while others are not. Two kinds of approaches dominate this literature: case studies of regions that are identified to be economically successful and more general theoretical approaches that aim to identify some of the circumstances that cause regions to be successful.

Using one or the other of these kinds of approaches, researchers have identified quite a few mechanisms and circumstances that seem to be responsible for, or at least influence, the economic success of a region. Several theoretical concepts have been established that deal with those mechanisms and circumstances. However, each of these concepts focuses on different aspects of the local system with some of them based on specific schools of economic thinking. A general analysis that tries to comprise the different theoretical concepts is missing. This paper tries to make progress in this direction.

In the literature some papers can be found that compare the different concepts, like industrial districts, innovative milieux and industrial clusters, and discuss their differences (cf. van Dijk 1995, Markusen 1996, and Mailat 1998). However, these comparisons are based on characteristics of the local systems, like the size of firms, the characteristics of their relations, or local historical and cultural specificities.

The approach that is proposed here, by contrast, takes an evolutionary perspective. In some specific location at some time, an evolutionary process creates particular circumstances and dynamics that lead to a phenomenon which is called industrial district, cluster or innovative milieu in the literature. The aim is to understand why, when and where industrial districts, clusters or milieux evolve. This contrasts with the theoretical literature that is more concerned with the reasons for the economic success of already existing industrial districts, clusters or milieux. As an implication, a classification of industrial clusters and the likes has to be based on the processes that lead to their creation rather than on the final characteristics.

The approach suggested here shares the interest in the dynamics that lead to the evolution of localised industrial clusters with most of the case studies. However, unlike those case studies which look into specific historical developments, it treats the underlying mechanisms in successful regions on an abstract, general level (as many theoretical works do without, of course, addressing the dynamic features).

The paper proceeds as follows. A definition of the phenomenon that is to be analysed in this paper is given in the next section. Section 3 contains a discussion of self-augmenting dynamics and the life cycle of localised industrial clusters. In Section 4, the basic elements of a regional system and their interactions are described. This leads to a classification of the self-augmenting processes that are involved in the evolution of industrial clusters which is given in Section 5. This classification is compared to the existing concepts in Section 6. Section 7 contains the classification of some regions for which case studies exist, according to the identified self-augmenting processes. Section 8 concludes.
2. Self-organisation and localised industrial clusters

This paper aims to improve the understanding of why, when and where industrial clusters evolve through an identification of the mechanisms that lead to their creation. Before the paper may proceed in this direction, it has to be defined what kind of phenomenon is intended to be studied.

In the literature, various definitions of local systems like industrial districts, clusters and innovative milieux are given and several typologies are proposed (see e.g. Aydakot 1988 and Markusen 1996). These typologies focus on the characteristics of existing clusters. Instead, the approach that is taken here focuses on the evolution of localised industrial clusters. The dynamics and mechanisms that lead to the existence of localised industrial clusters are of main interest. Therefore, the definition of localised industrial clusters that are studied here has to focus on different aspects than the definitions proposed in the literature.

Industrial districts, clusters and innovative milieux are usually defined as a local system of several interacting firms that belong to one or a few industries, and the respective surroundings, including aspects like the local education system, public research institutions, local culture and many more. The different definitions that can be found make various statements about the specific characteristics of these firms, their surrounding and the interactions between these elements. For example, in industrial districts, the firms have to be of small or medium size and their interaction has to be, according to many papers on the districts in North Italy, characterised by co-operation, competition and a high degree of division of labour. In the case of innovative milieux, instead, the definition also includes characteristics that describe the local institutions and public opinion, while the concept of clusters usually contains specific requirements with respect to the relations between firms. A common notion of all these concepts is the assumption that some local interactions exist between firms and other local elements that improve the competitiveness of the firms. Such an assumption is necessary since all these concepts aim to explain the success of the respective local systems.

This general assumption constitutes the central notion of the approach that is proposed here. In this approach, local systems are studied in which a number of firms of one or a few related industries mutually profit from their proximity and the local availability of other aspects that are due to their own existence. To this extent, it is in line with the concepts in the literature. However, it departs from them with respect to the local processes that are considered. Since it is concerned with the evolution of local systems, those local interactions that cause their existence are of interest, while in the literature, those mechanisms are studied that make local systems successful. There might be mechanisms that are important in both respects, but there also might be mechanisms that matter only in one of the contexts.

Thus, the basic characteristic of the systems that will be studied here is that a number of firms of one or a few related industries interact with a comparably high level of employment and themselves, and at the same time profits from this local system and shapes it. This means that there is a circular reinforcing mechanism in which the supportive local circumstances keep the employment in the firms high while at the same time this high employment in the industry or industries causes the supportive local circumstances. Taking an abstractive perspective, this means that there are aspects that positively influence each other and therefore keep their state on a level that is far
above the state that would be expected without such positive influences. This characterises a self-organising system.

Abstracting even further, the local system might be reduced to two elements, for example the total employment in an industry and the total amount of local support for firms within this industry. Of course, such an abstraction is far away from reality, but it helps to illustrate the structure of the mechanisms that are studied here. Below, the local processes will be analysed in more detail. If there are two local elements that positively influence one another as soon as they have reached a certain level, a system is constituted that has at least two stable states. As long as both element are in a poor state, their interaction does not help them to leave this state. If, instead, both elements have reached a sufficiently high state, their mutual support drives them towards even higher states until they reach some natural limit. For simplicity, let us assume that these two states are the only stable states in the system. This means that endogenous dynamics exist that move the local system always to one of the two stable states. Which of the two states is reached depends on the initial state.

Although the systems that are studied here are much more complex, it seems to be adequate to assume that two possible states exist, one with a high level of employment in one or a few industries in a region and one with a low level of employment. Of course, there are industries that do not show such a feature and one of the aims of the approach that is proposed here is to identify characteristics of industries that are responsible for the fact that such a feature appears. In this paper, however, it is assumed that such a feature is given. Hence, this paper only applies to industries with the respective characteristics.

Given a situation in which the local system converges with respect to a certain industry into one of two stable states — a state with a rather low activity and employment in the region and a state which is called a localised industrial cluster here — the evolution of a localised industrial cluster can be understood in the following way. The initial situation is one in which the economic activity with respect to the considered industry or industries is quite low in the region. This means that there is no firm or only a small number of firms with rather low numbers of employees located in the area and that there are no specific circumstances that would cause a significant improvement in the situation, always with respect to the considered industry or industries. The endogenous dynamics, meaning the mechanisms within the local system, keep the system in the low level stable state. This holds for constant exogenous conditions. A localised industrial cluster might evolve if the exogenous conditions change. A change in the exogenous conditions might have two kinds of effect on the local system.

First, it might change the strength or structure of the endogenous mechanisms that are responsible for the existence of two stable states. For example, a change in the national education system might influence the local accumulation of human capital and therefore enable a region to develop into a localised industrial cluster. A change in the endogenous processes, due to exogenous changes, usually lead to a movement in the critical state. The critical state, often called critical mass, is the state that separates those initial states that lead to a convergence into the low level stable state from those initial states that lead to a convergence into the high level stable state. If the convergence process within a region is comparably slow, a sudden change of the critical mass might cause the actual state to be above the critical mass, although it was the low level stable state before the exogenous
change. In such a case the endogenous processes will then move the local system towards the high level stable state.

Second, a change in the exogenous conditions might have a direct impact on the state of a local system. For example, a sudden increase in the demand for a certain good will cause the existing firms to increase in size until the new demand is satisfied again. In such a way the employment within a region might increase significantly without any change in the endogenous mechanisms. This might move the state of the local system above the critical mass. As a consequence, when the exogenous impulse due to unsatisfied demand vanishes, the local system will converge into the high level stable state due to the endogenous self-organisation.

These are the phenomenon that are to be studied here. Therefore, a localised industrial cluster is defined here as an agglomeration of firms of one or a few related industries in an area which satisfies two additional conditions: first, the firms have to profit mutually from the proximity of each other and the local circumstances that are created by the agglomeration and second, the agglomeration has evolved due to a positive feedback between geographical proximity and supportive interaction of the firms and other local aspects.

3. The life cycle of industrial clusters

To understand the evolution of such a localised industrial cluster, it has to be studied how this combination of supportive interaction and geographic proximity develops. Using the above definition of localised industrial clusters, this question can be reformulated into the question of how a bunch of firms that are active in one or a few industries come to be located in the same area and interact with each other or with their environment such that they mutually profit. A life cycle concept for localised industrial clusters can be developed according to the above theoretical considerations. If industrial clusters evolve due to the mutual reinforcement of geographic proximity and supportive interaction between firms, the constitutive process looks similar to the so-called lock-in processes that are comprehensively studied in the literature (cf. Arthur 1989 and Witt 1997). The starting point is a region at the low level stable state. Without any exogenous changes, the local system will remain in this state for ever. Therefore, the first phase of the life cycle of a localised industrial cluster encapsulates the change in the exogenous conditions and their impact on the local system. If, due to these exogenous changes, the critical mass is passed, the endogenous self-organisation leads to the development of the local system into a localised industrial cluster. This is called the second phase. Once the high level stable state is reached and the exogenous conditions remain fairly constant, the localised industrial cluster prevails. This is called the third phase. Finally, in the fourth phase, external changes might reduce the localised industrial cluster to a state below the critical mass, so that it disappears. Changing market conditions are a typical example of such external changes. Often the advantages that have been created in a region during the existence of a localised industrial cluster can be used for the establishment of a cluster in another industry. Thus, a region does not necessarily become a backward region if an industrial cluster disappears. However, considering localised industrial clusters, with their connection to one or a few industries, these four phases describe the typical, although not necessary, life cycle of these clusters.
Case studies support such a view. Especially the first two phases are extensively described in these studies.

A study of the evolution of localised industrial clusters has to focus on the first two phases of this life cycle. According to the above theory, two different kinds of aspects play a role. First, some (exogenous) initiation is needed to move the state of the local system above the critical mass. Second, there have to be mechanisms that constitute the self-organisation process which is responsible for the existence of two stable states. These mechanisms have to be self-augmenting so that the localised industrial cluster evolves endogenously, once the critical mass has been passed. If no such endogenous mechanisms exist, there is no multiplicity of stable states and the region does by and large only react to the exogenous conditions. What is called a localised industrial cluster here cannot evolve in such a situation. Hence, two kinds of prerequisites have to be given for a localised industrial cluster to evolve: there have to exist some initiating forces and some self-augmenting processes. To understand the evolution of localised industrial clusters, both prerequisites have to be understood comprehensively. While the specific local circumstances might provide the initiation for the evolution of an industrial cluster, the internal self-augmenting mechanisms determine the evolution itself (the importance of the initial dynamics have been discussed with the help of simulations in Brenner & Weigelt 2001). Thus, the internal self-augmenting mechanisms also determine the value of the critical mass and the local aspects that are important for reaching this critical mass. Thus, only if we understand the self-augmenting mechanisms and their impact on the dynamics of the local system is an analysis of the requirements for passing the critical mass is feasible. Therefore, this approach will focus on the internal self-augmenting mechanisms. It constitutes the first step in a study of the evolution of localised industrial clusters. Further steps have to follow.

We might take the argument one stage further. If the self-augmenting mechanisms exist to a sufficient extent for an industry, there will occur by chance in some regions the necessary conditions for a localised industrial cluster to evolve. Thus, if we restrict the analysis to the questions of why and when localised industrial clusters evolve, an understanding of the endogenous self-augmenting mechanisms provides valuable answers. Only if such local mechanisms exist for an industry, respective localised industrial clusters might evolve. Thus, new localised industrial clusters evolve when new industries with the respective characteristics come into existence or when the characteristics of an industry change in a certain way. Localised industrial clusters then evolve due to the local self-augmenting mechanisms and some random events that provide the initiation. If the question of where localised industrial clusters evolve is to be answered, the initiation processes and local circumstances have to be studied in more detail.

4. Elements of and mechanisms within localised industrial clusters

The aim of this approach is to identify the internal mechanisms that cause the evolution of localised industrial clusters, and then to study their influence on the evolution of existing industrial clusters. To this end, three steps are necessary: the identification of all interactions and dynamics within the local industrial system, the detection of all mechanisms that might constitute a self-augmenting process, and the analysis of these mechanisms in existing localised industrial clusters. The first step is done in this section.
Above, industrial clusters have been defined as a local system of firms. Thus, firms will be the central elements in the following analysis. In industrial clusters, their relation is characterised by two aspects: their geographic proximity and their supportive direct or indirect interactions. The mechanisms that are to be identified in this approach are those mechanisms that relate these two aspects to each other. Furthermore, this approach is restricted to the evolution of localised industrial clusters - the second phase according to the above described life cycle - caused by a simultaneous growth of geographic proximity and supportive interaction. Such an evolution goes together with an increase in the number of interacting firms or their size.

Therefore, the search for self-augmenting processes might take the increase in the number of firms and the increase in their size as the central element of analysis. All aspects that cause such processes and all local circumstances that profit from such processes have to be studied. An increase in the number of firms is reached either by the occurrence of start-ups or by the set-up of branches in the region by larger firms. Movement of firms is seldom and therefore neglected here. Firms generally grow if they have a competitive advantage over other firms or if they supply a growing market. The latter reason is caused by external circumstances, so that this approach focuses on the first reason, the competitive advantage. Hence, the following analysis concentrates on three processes, namely start-ups, set-ups of branches and the increase in the competitive advantages of firms.

To obtain a virtuous cycle, there has to be a local aspect that simultaneously profits from one of the above aspects while causing them. It has to be an aspect that also changes on a similar time scale. Aspects like culture and historical circumstances, which are repeatedly mentioned in the literature, are not relevant here since they do not change in the short run (they might, however, be important in the first phase of the life cycle of a localised industrial cluster). Six such elements can be found: other firms, the labour market, the capital market, academic institutes, public opinion and politics. The detailed interactions are discussed below. It should be kept in mind that only those interactions are relevant which include geographic proximity as an important aspect.

Other firms: Firms located in the same region influence each other in many way. A distinction between direct and indirect influence is helpful for structuring the mutual influence between firms. Indirect influence includes all those mechanisms where one firm changes the circumstances for another firm. Examples are the creation of human capital or a supportive infrastructure in the region. This indirect influence will be discussed in detail later. First, direct influence will be studied. There are four relevant direct influences: information flows due to informal contacts between employees of different firms, information flows due to formal co-operations, mutual profits due to joint projects and supplier-buyer relations.

The first three of these influences are empirically difficult to separate. So far, there exists no study that is able to identify the sole effects of one of these mechanisms. There is, however, quite a literature on knowledge and technological spillovers. This literature studies the influence of research in one firm or industry on the innovativeness of other firms or industries. There is strong empirical evidence for such an influence (cf. Grupp 1996 and Audretsch 1998) and that this influence depends on the geographic proximity of the involved firms (cf. Jaffe, Trajtenberg & Henderson 1993, Anselin, Varga & Acs 1997 and Maunseth & Vespagen 1998). Within a region spillovers occur within firms of the same industry,
between firms of different industries and from university and public research institutes to firms. R&D activities in firms, instead, seem to have no significant impact on the number of innovations by universities and research institutes.

However, it is not clear what these studies measure. These spillovers might be the result of informal contacts between employees of different firms, of co-operations between firms, or of flows of the working force (for a theoretical discussion of the different explanations of spillovers see Camagni 1991, Branstetter 1998 and Dahlum, Holmen, Jacobsson, Praest, Rickne & Villumsen 1999). Empirical evidence for each of the mechanisms separately is rather seldom.

There are some hints for the importance of information flows due to informal contacts (see e.g. the case studies in Kozul-Wright 1994 and Brown & Hendry 1998). The theoretical argument runs as follows. The proximity of firms causes their employees to meet privately in bars or through other social activities. In these informal contacts, information about job-relevant developments are exchanged (see Camagni 1991 for a theoretical discussion). This information flow increases the innovativeness of the firms and, even more so, the diffusion of innovations. Whether this mechanism works in all industries is an open question.

Similarly, it is not clear how much co-operation between firms contributes to the performance of these firms or in which way such co-operation influences the competitiveness of each of the firms (cf. the different findings in Malmberg 1996, Staber 1996, Grotz & Braun 1997, Koschatzky 1998, Oerlemans, Meeus & Boekema 1998, Panciota 1998, Fritsch 1999 and Harabi 1999). There is a clear connection between the innovativeness of firms and their contacts to other firms (cf. MacPherson 1997). Furthermore, co-operation, especially with suppliers and customers, is frequent and regarded as profitable (cf. Sengenberger & Pyke 1992, Vou & Schmitz 1992, Camagni 1993, Wilkinson 1994, van Dijk 1995 and Vipraio 1996). Thus, co-operation seems to have a positive effect on firms. Whether this positive effect is caused by flows of information or by the higher innovativeness due to complementary competences is less clear. The role of geographic proximity is also not clear. Co-operation appears mainly between firms that are located in the same region (cf. Koschatzky 1998, Oerlemans, Meeus & Boekema 1998, Sternberg 1998 and Fritsch 1999). However, the intensity of the formal contact between firms seems not to depend on the geographic distance between them, once the contact is established (cf. Hahn & Gaiser 1994 and Oerlemans, Meeus & Boekema 1998). Hence, it is not known whether having a higher number of potential co-operation partners nearby does enhance the co-operative activities of a firm. Therefore, the assumption that a concentration of firms of the same industry enhances co-operation between them, so that they profit from this concentration, remains speculative.

Besides the above discussed co-operation, there are some specific situations where co-operation between firms is necessary for their survival. If this repeatedly occurs, there is a mutual dependence between firms. An example is the film industry (see Enright 1995). A product cannot be developed there without co-operating with other firms. If, in addition, proximity makes co-operation easier, an increase in the number of firms in one region increases their possibilities for co-operation and therefore improves their competitiveness. However, except for the film industry, such a situation has not been reported for any other industry, although it might be claimed that co-operation creates mutual dependence (see Arrighetti, Bachmann & Deakin 1997 for such an argument).

When Marshall (1947) discussed the advantages of industrial districts, the proximity of
suppliers and customers has been an important aspect. In the past decade empirical studies have shown that the proximity of suppliers and customers has become less important (cf. e.g. Hahn & Gaiser 1994, Grotz & Braun 1997), mainly due to the strong decrease in the relative costs of transportation. What remains is a slightly higher intensity of contacts if supplier and customer are located near to each other (cf. Fritsch 1999). Whether this improves the competitiveness of the firms and therefore leads to growth cannot be deduced from the empirical studies. In the case of start-ups and the set-up of branches, some higher activity can be observed in regions where customers are located. Existing firms in a region seem to induce especially the founding of start-ups that supply these firms in the same region. For customer firms no such effect can be observed. Therefore, no clear virtuous circle is established. There is only a positive feedback from customers to suppliers. There is no positive influence from suppliers on the set-up of customer firms, and there is no empirical support for a positive influence of suppliers on the growth of customer firms.

Labour market: The local labour market has a positive impact on all three processes: the set-up of branches, start-ups and the competitive advantage of firms. It is usually classified as one of the most important criterion for the choice of the location of branches. Employees are in general not willing to move, so that firms mainly rely on the local labour market. In the literature on industrial districts and innovative milieus, similar arguments are put forward. There the existence of an adequate labour force is also seen as a crucial factor for the success of such regional systems (cf. Maillat & Lecoq 1992, Zeitlin 1992 and Pietrobelli 1998). However, this approach goes one step further and distinguishes between two kinds of human capital: transferable and non-transferable, or specific human capital (for a similar distinction see Huiban & Boushina 1998 or Becker 1993). Transferable human capital is based on academic and scientific knowledge and can be obtained in schools and universities, while non-transferable human capital is accumulated on the job. The need for this kind of human capital differs between firms. Some rely mainly on transferable human capital, some on non-transferable human capital, and some firms require a large amount of both.

The above argument implies that adequate human capital in a region improves the competitiveness of the firms in this region. Thus, the same mechanisms hold for the influence of the labour market on the competitiveness of firms as on the set-up of branches.

The literature on start-ups reveals that most start-ups in the manufacturing sector are founded either by researchers that worked before in universities or research institutes, or by employees of firms that operate in a similar product market or with similar technologies. Furthermore, most start-ups are located in the region or near the region where its founder was working or living before (cf. Bramanti & Senn 1990 and Dahlin, Holmen, Jacobsson, Praeis, Rickne & Villumsen 1999). Thus, the number of start-ups should be expected to depend strongly on the human capital and the number of related firms in the region (see Audretsch & Fritsch 1999 for some empirical evidence). Transferable and non-transferable human capital plays a role for start-ups.

In addition, the number and size of firms in a region has also an influence on the labour market there. Workers usually improve their skills on the job. Thus, the human capital in a region increases due to the economic activity there. However, although firms also use seminars to train their employees theoretically, the main impact of firms on the stock of
human capital relates to non-transferable human capital. During work employees collect a lot of experience. If they leave a firm, they take this tacit knowledge with them and enrich the labour market (a detailed study of such a process is given in Tomlinson 1999). Of course, their tacit knowledge is related to their previous job, which often implies that the use of this knowledge is restricted to an industry or even a technology. It might be useful only to a certain extent for other firms. However, many firms rely on such a kind of human capital (cf. empirical evidence in Bramanti & Senn 1990). Hence, a higher number of firms in a region generally implies a higher amount of human capital there, especially non-transferable human capital.

Local capital market: Investment flows have become increasingly global in the last decades. Therefore, the local capital market is not important for those firms that have access to the global capital market. Those are the larger and well-established firms. Start-ups and small firms, instead, lack access to the global capital market and rely heavily on the local capital supply. Start-ups profit a lot from local venture capitalists, not only in terms of the provision of necessary capital but often also in terms of the business experience that venture capitalists have (see Maillat & Lecoq 1992 and Rickne 2000). Thus, there is clear evidence that the local availability of venture capital and experienced venture capitalists increases the number of start-ups and the likelihood of their success.

Whether there is also a causation that runs in the opposite direction is less clear. Local firms often provide the capital for start-ups, especially for start-ups that operate in related areas or are even suppliers (in Rickne 2000, they have been found to be involved in 22% of initial financing and 67% of development financing). This implies that a higher number of firms in a region might improve the local availability of venture capital. In addition, local banks play an important role in the supply of capital for small firms. Local banks have been found to be more willing to lend money to firms that operate in industrial or technological fields where these banks have already collected experience in the past. Therefore, the existence of quite a number of firms in a region should imply that the banks there are experienced in the respective industrial or technological field and should be more willing to provide capital for similar ventures. This line of argument has not been tested empirically. Thus, there is no clear empirical evidence for the proposition that a high number of firms of one industry increase the availability of venture capital in the respective region. Nevertheless, there seem to be some findings and considerations that support this proposition.

Public research institutes and universities: Universities and research institutes influence all three kinds of processes. Quite a number of start-ups are founded by people who worked in research institutes or universities before. Thus, academic institutes are one of the main sources of foundens in a region. Furthermore, they are a main source of qualified labour. On the one hand, universities offer the opportunity to study. Through this they create transferable human capital. On the other hand, people who do applied research in universities and research institutes obtain experience which they take with them if they get a job in a firm or start their own firm. Thus, universities and research institutes also create non-transferable human capital, although to a smaller extent than firms and mainly restricted to tasks that are connected with research and development. People who have done applied
research in universities or research institutes usually possess a high amount of transferable and of non-transferable knowledge. This human capital is an important reason for firms to set-up branches in certain regions, and it is a necessary ingredient for firms to grow. Finally, research institutes that engage in applied research offer the possibility to conduct joint research projects and are a major source of expertise for firms (cf. MacPherson 1997). Firms in the region can profit more than other firms from such joint research projects and might obtain comparative advantages.

However, a causation in the opposite direction is less evident. An increasing number of firms in a region does not automatically imply an increasing number of public research institutes and universities. Sometimes new institutes are founded due to the strength of a certain industry in the region. In general, however, the establishment of public research institutes and universities is a political decision. The influence of firms on such decisions is rather weak. In addition, firms have some influence on the fields of research in public research institutes, mainly due to joint projects. Nevertheless, the overall influence of an increasing number of firms in a region or their growth has little impact on the number and orientation of public research institutes and universities there. There is also no positive impact found of the firms' activities in the region on the output of research in public institutes (cf. Anselin, Varga & Acs 1997).

Public opinion: It is sometimes claimed in the literature that public opinion plays an important role in industrial districts and innovative milieux (cf. Miller & Côté 1985). Empirical evidence for this claim, however, is weak. There is some evidence for the claim that the common beliefs about the economic perspectives of a region or industry influence the number of start-ups. Besides this, there is no convincing explanation of why public opinion should increase the number of start-ups, the set-up of branches or the growth of firms in a region. Hence, the proposition that public opinion in a region might enhance economic activity there can only be supported by the convincing, but not empirical tested, argument that positive expectation about the future of a region and the identification with an industry makes start-ups more likely to occur.

A causation in the opposite direction has to rely on a similarly convincing, but not empirical tested, argument. Each new successful firm and the growth of firms should increase the faith in a positive economic development of the region and industry. Thus, there seems to exist a virtuous cycle which, however, has not been proved empirically.

Local policy: Policy has a strong influence on the local circumstances. It determines the founding and orientation of universities and public research institutes, the infrastructure, and the specific conditions for firms. The political situation in a region, including the tax and the education system, are also an important factor for the choice of location (e.g. in the case of the radiocommunications cluster in Northern Denmark as described by Dalum 1995). Furthermore, in many countries local government activities can be found that support start-ups (cf. Miller & Côté 1985). Thus, policy influences the number of start-ups, the set-up of branches, and the growth of firms in a region.

The impact of firms on policy, instead, is restricted. The engagement of governments in regions is restricted by budget constraints. Some fundamental aspects of policy are bounded by the general public opinion in the country which often does not change quickly. Thus,
politicians are only able to react to the claims of firms up to a point. Most of the shifts in policy are shifts in priorities. These, however, can be influenced by firms or their lobbying institutions. The more firms and employment in a region are related to one industry, the more influence will this industry have on the political priorities in the region (see Bramanti & Senn 1990 for some evidence for the influence of firms on the regional education system). Therefore, although there is no comprehensive empirical evidence for this, an increasing number of firms or the growth of firms in a region will in general lead to a policy that favour these firms more.

5. Classification of self-augmenting forces

Above, a distinction has been made between the exogenous circumstances that may serve as prerequisites for the evolution of industrial clusters and the internal interactions that cause such an evolution (cf. a similar argumentation in Garnsey 1998 and Garnsey & Smith 1998). In order to understand why, when and where industrial districts evolve, both aspects have to be studied in detail. However, the study of the necessary exogenous circumstances requires an understanding of the self-augmenting internal interactions. Therefore, this approach provides a comprehensive analysis of local self-augmenting processes.

In the previous section all the possible local mechanisms that might lead to self-augmenting processes have been discussed. The empirical evidence for each of these processes has been outlined. As a consequence, several positive feedback-loops can be identified. There are seven such virtuous cycles that will be discussed separately below.

Accumulation of human capital: The above discussion has shown that there is empirical evidence for the dependence of firms on the local labour market and the human capital that is embodied in this labour market. Similarly, firms create human capital due to the collection of experience on the job. A positive feedback-loop is created as a result of this. The development of human capital in one firm can be used by other firms – due to spin-offs and the movement of employees – which then, in turn, if they develop will also supply non-transferable human capital to the local labour market. This process requires that the non-transferable human capital that is required and created by different firms in the region has to be similar enough so that these firms can profit from each other.

The two processes, however, take place on different time scales. Firms immediately profit from the existence of an adequate local labour market and the number of start-ups is high if the respective local human capital is high. The creation of human capital within firms, instead, takes some time. Thus, the virtuous cycle that is created by these mechanisms can be expected to work slowly.

There are two areas of economic activity in which non-transferable human capital (in the literature authors often talk about tacit knowledge; see e.g. Antonelli 1999 and Maskell & Malmberg 1999) is especially important. These are research and development on the one hand and craft-based production on the other hand, including knowledge about the kind of products that will be demanded. The accumulation of human capital in these two areas can be easily distinguished in case studies. The evolution of existing industrial clusters has usually been supported by the accumulation of one of these kinds of human capital. Furthermore, the necessary circumstances for these two processes of human capital accu-
mulation differ. Therefore, it seems to be helpful to distinguish between the two respective self-augmenting processes.

Furthermore, as has been argued above, the availability of human capital in a region has an impact on several aspects of the population of firms in the region. First, a highly skilled labour force is in general more productive, so that firms in the region profit from the human capital in the form of higher productivity. This is a competitive advantage and should help the firms to grow. Second, experience that is collected on the job is often the basis for founding an own firm. Most start-ups are spin-offs founded by former workers of firms in the same region. Thus, a high human capital increases the number of start-ups in a region. Third, qualified labour is crucial for the development and invention of new products and processes. Firms that can rely on a high human capital are more innovative and therefore more competitive. Again, a growth in these firms should be the result. Each of these three mechanisms in combination with the fact that firms create human capital constitutes a virtuous cycle.

Informal contacts causing information flows: Above, it has been argued that the proximity of firms enhances informal contacts between their employees. This leads to an information flow that increases the innovativeness of both firms and reduces the time they need to adopt innovations by others. They therefore become more competitive and grow. This, in turn, increases the possibilities of informal contacts. There is some empirical evidence for such a self-augmenting process. However, informal contacts are difficult to study and seldom addressed in the empirical literature. Therefore, it is not known whether this mechanism works in all industries.

If it works, it can be expected to do so immediately. The informal contacts between employees often result from former formal relations; for example, studying at the same university. In contrast to the self-augmenting process based on the development of human capital, the information flow through informal contacts only influences the competitiveness of firms. Thus, it does not directly increase the number of firms in a region but might increase their likelihood to survive and their speed of growth.

Mutual dependence of firms: There might be cases where firms can only survive if they co-operate or co-exist with others. In this case they mutually depend on each other. An example for such a mutual dependence is the film industry where nowadays a company is usually not able to produce a film on its own. The necessary co-operation with other firms is more easily reached if the firms are located nearby. The mutual dependence motivates firms to move into the same region. In turn, the location in a region where other firms are located increases the likelihood of survival. A self-augmenting process is created. In other industries, such a clear mutual dependence has not been found. One might consider hub-spoke districts as a similar case. There one large firm is located in a region together with a large number of its suppliers. Generally, the suppliers depend on the buyer firm. But this does not constitute a mutual dependence since the buyer does not depend completely on the suppliers.

The mutual dependence of firms leads to two processes. First, it attracts firms to regions where other firms are already located. Second, it increases the competitiveness of the firms in such regions.
Co-operation between firms: It has been above reported that most co-operation between firms are local. Thus, one might claim that the more firms with related or complementary interests are located in a region, the more options for co-operation exist in this region. This might increase the competitiveness of these firms. As a consequence they can be expected to grow and the options for co-operation within the region increase further. This would constitute a self-augmenting process. However, the empirical literature does not answer the question of whether it is more profitable for firms to have co-operation partners that are located in the same region than to have to search for co-operation partners outside the region. The fact that most co-operations are local might be caused by the fact that firms start searching in the local environment and that they are better informed about firms in the same region. Co-operation also occurs between firms that are located far away from each other. Thus, the lack of local co-operation partners might not restrict the competitiveness of firms.

Nevertheless, it seem plausible that firms somehow profit from having co-operation partners nearby; for example, due to small transaction costs. Thus, it might be assumed that firms receive some competitive advantage from the fact that they are located in a region with many other firms that operate in the same or a related field. The profit can be either an increase in productivity, for example, if the firms jointly use some production devises, or a higher innovativeness, for example, due to joint R&D projects. Both lead to a higher competitiveness on the market and therefore to a higher chance of growing. These mechanisms need some time to take effect. Thus, together with the higher potential for co-operation once the firms have grown, a virtuous cycle is created that slowly increases the economic activity in the region.

However, another virtuous cycle can be identified in connection with co-operation between firms. The number of co-operations within a region does not only depend on the number and size of the firms that are located there. It also depends on the attitude of the respective persons in the firms towards co-operation. Besides cultural influence, the experiences in the past significantly influence this attitude. Although there is again no clear empirical evidence, it seems to be plausible to assume that successful co-operation increases the likelihood of further co-operation. The same might be reached by the transmission of information about successful co-operations in the region. Thus, the number of co-operations and the attitude towards co-operations might as well constitute a virtuous cycle that contributes to the profitability of the firms in the region.

Local capital market: Empirical studies show that start-ups rely strongly on local sources of finance. Part of this finance comes from venture capitalists, another part comes from other firms operating in the same industry. Thus, if firms of a certain industry already exist in a region, they increase the probability for a start-up of the same industry to receive venture capital. However, there is no empirical evidence so far that this holds for all industries. It is also often claimed that venture capitalists and banks tend to invest in industries in which they obtained experience in the past. This would again favour start-ups in regions where many other firms of the same industry are located. Both effects constitute a positive feedback-loop. However, the empirical evidence is not sufficient to claim such a self-augmenting process for all industries.

If such a self-augmenting process exists, it should lead to an increasing number of start-
ups with some time delay. Local venture capitalists and banks need some experience before they will start to intensify their activity in a certain industry. Similarly, firms have to grow before they are able to provide further start-ups with venture capital.

Public opinion: There is some theoretical discussion about the influence of the attitude of the local population on the development of industrial clusters. This is based mainly on the identification of the population with a certain industry and the beliefs about a positive future development of the region. Both aspects might constitute a self-augmenting process. Empirical evidence, however, is missing. Nevertheless, many researchers share the experience that in successful regions the people have on average a positive expectation about the future of their region while in backward areas they tend to be pessimistic. The direction of the causation is not that clear. However, it seems to be plausible that the opinion of the population somehow adapts to the real situation. At the same time, the opinion on future developments is crucial for the decision to start an enterprise or expand into new markets. Furthermore, the expectation with respect to a certain industry also influences the decision of people to invest in the accumulation of respective industry-specific human capital.

Given that these processes exist, they would mainly influence the number of start-up in a region and to some extent the competitiveness of the existing firms due to the fact that they have better chances of attracting people. Both mechanisms take some time to take effect because public opinion adapts slowly to changing circumstances.

Local policy: Policy has a strong influence on the number of start-ups in a region, the movement of firms to a region, and the competitiveness of firms in a region. Whether, in turn, the existing firms have an influence on political decisions in such a way that a self-augmenting process is established is less clear. It might be assumed that the more firms with similar interests are located in a region, the stronger is their influence on policy. Such a claim seems to be plausible, but whether this leads to a higher number of firms in the region or to a competitive advantage of the firms is not clear.

There might be a mechanism that leads to a permanent improvement of the local conditions for firms of a certain industry and an increasing influence of this industry on policy. Such a process can be expected to work continuously. However, its existence is not guaranteed.

6. Discussion of the existing concepts

Above, seven classes of self-augmenting processes have been identified, which are claimed to be responsible for the evolution of industrial clusters. In the literature many case studies and several theoretical concepts can be found. Thus, the final step in this approach is a discussion of the existing concepts and case studies in the light of these self-augmenting processes.

In the literature four concepts of local systems are repeatedly used that are of interest here. These are ‘industrial districts’, ‘technological districts’, ‘innovative milieux’ and ‘economic clusters’. Each of the four concepts will be discussed in detail. It will be analysed which of the above self-augmenting processes play a crucial role within these concepts. This comparison faces one specific difficulty: while in the literature the characteristics of existing localised industrial clusters are studied, the mechanisms that led to their evolution are
considered here. As a consequence, several aspects that are regarded as important characteristics in the literature are results or even optional results according to this approach. This should be kept in mind during the following comparison. The concept that is proposed here will be briefly reflected upon after the four concepts have been discussed. Finally, a comparison of the central aspects of all five concepts is given in Table 2. In this comparison only the local self-augmenting processes are considered. It becomes clear that the concepts from the literature focus each on some of the mechanisms that are identified here. This implies that in each of the real situations only some of the mechanisms are relevant. However, a general theory has to include all these mechanisms to be able to explain the different cases. This will become more apparent in the next section.

Industrial districts: In the literature several and not always the same requirements are given as a definition of industrial districts. The term goes back to Marshall who has identified three typical characteristics of such districts: the specialisation due to a division of labour, the local creation and availability of a specified labour force, and economies of information and communication (cf. the discussion in Zeitlin 1992). The first of these aspects has not been identified above as a source of self-augmenting processes. It is seen here as the usual result of the agglomeration of specific economic activities in a localised industrial cluster and not as a mechanism of its evolution. The second aspect has been studied above under the label of the accumulation of human capital in a region. It has been stated that there is strong evidence for the influence of such a process. The last aspect, the economies of information and communication, correspond in some way with the above-discussed informal contacts and the resulting information flows. There is some empirical evidence for such a process.

With the intensive studies of several regions in North Italy, the concept of industrial districts has been changed significantly (cf. Becattini 1990 who started this line of research). It has become narrower and the label ‘Italian industrial district’ is often used for such local systems. In addition to the aspects that have been put forward by Marshall, the existence of a network of entrepreneurs with similar cultural background (cf. Trigilia 1992, van Dijk 1995, Rabellotti 1997), the frequent co-operation between firms (cf. Sengenberger & Pyke 1992, Vou & Schmitz 1992, Wilkinson 1994, van Dijk 1995, Vipraio 1996), and a large number of small firms (cf. Schmitz 1992, Sengenberger & Pyke 1992, van Dijk 1995, Rabellotti 1997) have become central in the discussion on Italian industrial districts. These aspects are not seen as self-augmenting mechanisms here, except for the co-operation between firms. The cultural background seems to play an important role in triggering the evolution of industrial districts but does not constitute a self-augmenting process. The restriction to agglomerations of small firms highlights a feature of the Italian districts which might restrict a general analysis too much (cf. a similar argumentation in Scott 1992).

Technological districts: The concept of technological districts has much in common with the concept of Italian industrial districts. It is also based on a region with a large number of small enterprises that co-operate with each other (cf. Storper 1992 and Dahum 1995). However, the focus is more on the learning process and the persistent innovative activity in the region than on the craft-based production process. This corresponds quite well to the distinction between non-transferable human capital in the area of R&D and of craft-based
production that has been proposed for the self-augmenting process that is based on the accumulation of human capital. Thus, the concept of technological districts contains two of the above-discussed self-augmenting processes: the accumulation of human capital and co-operation between firms.

Innovative milieux: The concept of innovative milieux is based on the notion of a local learning and innovation process. Learning and innovations are not seen as individual processes but as the collective actions of a number of firms and actors that are connected by networks, informal contacts and a common identity. Also the local resources of human capital and their immobility are seen as important aspects (cf. Camagni 1995 and the comparison with an industrial district in Lawson 1997). Comparing the concept of innovative milieux with the classes of self-augmenting processes that have been identified above, three of these processes are included in the concept: the accumulation of human capital, information transfers through informal contacts, and the specific public opinion with respect to innovations and the region.

Economic clusters: The term ‘cluster’ has become frequently used in recent economic literature (a general approach can be found in Drejer, Kristensen & Laursen 1997). Expressions like ‘regional cluster’, ‘industrial cluster’ and ‘innovative cluster’ are common. All of them are defined by a network of firms that are connected with each other through supplier-buyer relations, knowledge transfers, co-operations or similar things. Furthermore, specific circumstances like a specific labour market, specific institutions or a specific infrastructure are seen as part of such an economic cluster. However, the central aspect is the network of directly connected firms. This aspect corresponds to the self-augmenting process of co-operation.

Localised industrial clusters: The concept that is proposed here contains all the local self-augmenting mechanisms that might contribute to the evolution of localised industrial clusters. These mechanisms have been identified according to theoretical considerations and empirical evidence in the previous section. Seven mechanisms have been found to be theoretically plausible. For one of these mechanisms, the accumulation of human capital, strong evidence exists. Furthermore, the three mechanisms that are based on information flows due to informal contacts, co-operation between firms, and venture capital are likely to be relevant for the evolution of localised industrial clusters. Less clear is the impact of the mechanisms that are based on public opinion and the local politics. Both mechanisms are plausible enough to have an impact, but empirical evidence is missing. The mechanism that is based on mutual dependence is very specific. It only occurs in certain industries.

7. Self-augmenting processes in case studies

After the analysis of the theoretical concepts, we will turn to some of the numerous case studies in the literature. For each of these case studies, those self-augmenting processes will be identified that play a crucial role for the evolution of the respective localised industrial cluster. This gives some further hints about the importance of the different mechanisms that have been identified above.
## Table 1: Comparison of the concepts repeatedly used in the literature with the concept that is proposed here with respect to the self-augmenting processes that are identified above.

<table>
<thead>
<tr>
<th>Localised industrial cluster</th>
<th>human capital</th>
<th>informal contacts</th>
<th>mutual dependence</th>
<th>co-operation</th>
<th>venture capital</th>
<th>public opinion</th>
<th>local policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marshallian industrial district</td>
<td>basic element</td>
<td>basic element</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Italian industrial district</td>
<td>included</td>
<td>included</td>
<td>-</td>
<td>included</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Technological district</td>
<td>included</td>
<td>included</td>
<td>-</td>
<td>included</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Industrial districts in Italy: The case studies of industrial districts in Italy repeatedly report that the rise of these districts was based on a large number of small start-ups, founded by people with the required crafting skills and an appetite for being entrepreneurs. Most of these start-ups had been founded in the backyard without much initial capital. The high human capital with respect to knowledge about crafting and trading played an important role (cf. e.g. Trigilia 1992 and Omiccioli 1999). Furthermore, informal contacts between the entrepreneurs and workers are often claimed to have been important in these regions, although empirical evidence is missing. For two other aspects that are often put forward there is some evidence. Policy has supported the evolution of these districts strongly (cf. Becattini 1997), although it is not clear whether the development of the districts also had an influence on the local policy there. Quite some evidence exists that the public opinion about entrepreneurship and the future of the region had quite an impact (cf. Becattini 1997). The success of some firms, in turn, supported the development of such positive public beliefs.

Silicon Valley: Another region that is usually seen as an industrial district is Silicon Valley (see Saxenian 1994). In this case, the localised industrial cluster was supported in its initial phase by a large number of orders from the government. While this is considered here as an external force, some internal mechanisms can be identified as well. A frequent switching of the labour force between the firms and many informal contacts between employees of different firms led to a high flow of knowledge and human capital within the region. In addition, the created human capital also led to a lot of start-ups. Therefore, two of the above
self-augmenting processes can be found in Silicon Valley: the accumulation of human capital and significant information flows due to informal contacts. These two seem to be, besides the external influence of the government, responsible for the evolution of this localised industrial cluster.

The radio-communication cluster in Northern Denmark In the last ten years a strong cluster of radio-communication firms has evolved in Northern Denmark (cf. Dahlum 1995). Most of the firms that are located there now do not produce in this area. Their activity is restricted to R&D departments which were established there because the necessary labour force could not be attracted to other locations but can be moved from other firms within the region. The specific human capital that is available in this region has been the main force to attract firms or their R&D departments to this region and to allow for successful spin-offs within the region. The creation of a related research institute at the university of Aalborg has started or helped to start this development as well as the specific market situation, but the self-augmenting process is based on the creation of specific human capital within the region.

The American film industry The American film industry is nowadays concentrated in Los Angeles. The reason for this concentration is that firms of the film industry are often unable to conduct projects on their own (cf. Enright 1995). Joint work is crucial for the success of projects and leads to some kind of mutual dependence between the firms, where firms may easily switch between partners. Co-operation is more easily established if the firms are located near to each other. This has led to a self-augmenting process as described above with the final result being a very concentrated film industry in the U.S.A. (cf. Enright 1995).

<table>
<thead>
<tr>
<th>districts</th>
<th>human capital</th>
<th>informal contacts</th>
<th>mutual dependence</th>
<th>co-operation</th>
<th>venture capital</th>
<th>public opinion</th>
<th>local policy</th>
</tr>
</thead>
<tbody>
<tr>
<td>in Italy</td>
<td>basic process</td>
<td>probably involved</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>involved</td>
<td>probably involved</td>
</tr>
<tr>
<td>Silicon Valley</td>
<td>basic process</td>
<td>basic process</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>North Denmark</td>
<td>basic process</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>American film industry</td>
<td>-</td>
<td>-</td>
<td>basic process</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2: Classes of self-augmenting processes that are involved in historical examples of industrial clusters.
8. Conclusions

This paper analyses the evolution of localised industrial clusters from a self-organisation perspective. Therefore, the internal self-augmenting processes are of core interest. Seven such processes have been identified. The paper aims at shifting the focus of research into industrial clusters to these self-augmenting processes and therefore to the evolution of industrial clusters. If we intend to understand why localised industrial clusters exist and when and where they evolve, we have to study the mechanisms that lead to their evolution. Hence, the empirical evidence for each of the possible mechanisms is discussed here in detail. For some processes clear evidence exists, while the effects of others has to be doubted. When screening the case studies on industrial clusters for self-augmenting process, a similar picture is obtained. The accumulation of human capital in the region plays a strong role in nearly all the analysed industrial clusters. There is also quite some evidence for the impact of information flows through informal contacts. In one specific case mutual dependence between firms plays a crucial role. Some evidence has been found for the impact of the interaction of firms with political and public opinion. The other mechanisms could not be supported by this study.

Hopefully, this analysis of self-augmenting processes in the context of industrial clusters helps to structure the theoretical discussion in some way and focus this discussion more on the processes that lead to the evolution of localised industrial districts. This would then allow us to understand in more detail why, when and where industrial districts occur and how this process can be politically influenced. Although the discussion about industrial districts started around 20 years ago and has obtained more and more momentum, it seems that there is still a lot of work to be done.

References


Rickne, A. (2000). Network-Based Innovation and Learning Regions - The Case of Young Biomaterials-Related Firms. mimeo, Chalmers University of Technology.


