Cluster Life Cycles – Dimensions and Rationales of Cluster Evolution

Article *in* Industrial and Corporate Change · February 2010

DOI: 10.1093/icc/dtp036 · Source: RePEc

2 authors:

Dirk Fornahl
Universität Bremen

Max-Peter Menzel
University of Bayreuth

Industrial and Corporate Change, Volume 19, Number 1, pp. 205–238
doi:10.1093/icc/dtp036
Advance Access published July 22, 2009

Cluster life cycles—dimensions and rationales of cluster evolution

Max-Peter Menzel and Dirk Fornahl

We present a model that explains how the very cluster dynamics is both the driver for the movement of a cluster through a life cycle and the reason why this movement differs from the industry life cycle. The model is based on two key processes: the first is that the emergence, growth, decline and renewal of the cluster depend on the technological heterogeneity of firms; the second is that firms have a larger relative absorptive capacity, when they are in the same location, and thus especially localized learning changes heterogeneity: it leads to a technological convergence when learning takes place within the cluster and technological divergence, when learning takes place outside the cluster, yet in the same region. We derive hypothesis from the model regarding different phases of the cluster life cycle.

1. Introduction

Clusters are regarded as important elements in economic development (Porter, 1998; Malmberg and Maskell, 2002; Martin and Sunley, 2003). Companies in clusters experience stronger growth and faster innovation than those outside clusters (Audretsch and Feldman, 1996; Baptista and Swann, 1998; Swann *et al.*, 1998). These characteristics cause clusters to be considered a prerequisite for regional prosperity (Bathelt, 2001; Porter, 2003). To understand the functioning of clusters, many factors that define clusters are detailed in the literature (see for example: Enright, 2003; Martin and Sunley, 2003; Sternberg and Litzenberger, 2004), followed by various theoretical contributions, ranging from the industrial district literature (Becattini, 2002) to Porter's (1990) Diamond and Malmberg and Maskell's (2002) work on a knowledge-based theory of the cluster. However, the strong research focus on the way clusters function is contrasted with a disregard for their evolutionary development, i.e. how clusters actually become clusters, how and why they decline and how they shift into new fields (see also Lorenzen, 2005; Boschma and Frenken, 2006). The few existing insights on the emergence of clusters, for example, lead to the conclusion that the processes responsible for the functioning of a cluster cannot explain its emergence (Bresnahan *et al.*, 2001; Orsenigo, 2001). In addition to this, examples of declining clusters (Grabher, 1993) illustrate that the economic advantages that stem from cluster dynamics are not permanent. In fact, the decline of clusters seems to be caused by factors that were advantages in the past (Jacobs, 1969; Martin and Sunley, 2006). These findings indicate two things: first, that theories that explain the dynamics of functioning clusters are not sufficient to explain their evolution. The second is that clusters follow a kind of life cycle with different phases or stages of emergence, growth, and decline that differ in their characteristics (Bergman, 2008).

At the first glance, it seems obvious that clusters follow the life cycle of their respective industry. Seen this way, the cluster life cycle is only the local expression of the superordinate industry. However, empirical studies indicate that the life cycles of clusters and industries are different. Different clusters that belong to the same industry life cycle can follow different growth paths, as prominently described in Saxenian's (1994) contrasting stories of the computer industry in Boston and Silicon Valley. Comparisons of clustered and non-clustered companies during the industry life cycle highlight additional differences: clustered companies outperform non-clustered companies at the beginning of the life cycle and have a worse performance at its end (Audretsch and Feldman, 1996; Pouder and St. John, 1996). This shows that the cluster life cycle is more than just a local representation of the industry life cycle and is prone to local peculiarities.

There are already some approaches that explain the movement of a cluster through a life cycle. Van Klink and De Langen (2001) describe the cluster cycle as a progression through the phases of development, expansion, maturation, and transition. Yet, it is unclear in their study, which influences stem from cluster dynamics and which from the industrial environment. Pouder and St. John (1996) argue that the movement through the life cycle is shaped by a biased cognitive focus of the cluster's companies toward each other. However, their cluster life cycle ultimately results in a negative lock-in and decline (Martin and Sunley, 2003), thereby neglecting the possibility of entering new growth phases. In the same vein, Maskell and Malmberg (2007) argue that myopic behavior of entrepreneurs reinforces existent patterns of specialization, which can both result in the emergence of a cluster and its negative lock-in. Yet, they give no indication of how a cluster can be renewed and why the cluster life cycle differs from the industry life cycle.

Currently, there is no approach that explains the emergence, growth, decline, as well as the renewal of clusters, and why the life cycles of clusters differ from those of industries. We intend to bridge this gap and present a model to analyzing the life cycles of clusters. The core premises of the model are that the movement of the cluster through the life cycle depends on the increase and decrease of heterogeneity among the cluster's companies and organizations; and that the way firms exploit this heterogeneity distinguishes clustered from non-clustered companies. In doing so, we move the analytical level from the cluster to the single firm and its relation to other firms. To give an outline of the model, we follow Malerba's (2006: 19) suggestion to "identify some empirical regularities, stylized facts or puzzles that need to be explained, develop appreciative theorizing, do quantitative analysis, and formal modeling". We concentrate on the first two elements of his suggestion below.

Section 2 summarizes stylized facts on different phases of clustering. In Section 3, we take one step back and begin with the question of what clusters actually are, to uncover the essential elements of cluster development. Section 4 analyzes the insights gained from research on industry life cycles and identifies analogies to the cluster life cycle. Section 5 presents the basic dimensions of the model and explores their interdependencies. Section 6 analyzes how localized learning results in a deviation of the cluster from the industry life cycle. We apply the findings to the different stages of the cluster in Section 7. Section 8 concludes the article and gives some suggestions for future research.

2. Stylized facts on cluster stages

The contributions that compare the development of clustered and non-clustered companies provide general evidence that companies in clusters grow more strongly and innovate more quickly than non-clustered companies and that clusters attract more start-ups than regions without a cluster (Swann *et al.*, 1998; Baptista, 2000; Klepper, 2007). But this picture changes when the development of clustered and non-clustered companies is compared over longer time periods or during different stages of the industry life cycle. At the very beginning of the industry life cycle, a distinct spatial concentration is not observable. Despite some small agglomerations, the small number of companies in the emerging industry is geographically dispersed (Klepper, 2007). Clusters begin to emerge as the industry grows. Klepper (2007), for example, shows for the automobile, tyre, and television industries that the phase of strong industry growth is accompanied by an increasing geographic concentration. Pervasive spin-off processes and higher growth rates of the clustered companies result in an increasing concentration of the whole industry. When the phase of strong growth ends (in the 1940s for the automobile industry, for example), the industry becomes more dispersed. He sees plants constructed in remote areas to avoid congestion effects and to move production closer to the intended market as responsible for it. In addition to this, the increasing codification of knowledge in the maturing industry decreases the necessity for a company to be near the places where this knowledge is generated. Ketelhohn (2006), applying the Ellison–Glaeser Index (Ellison and Glaeser, 1997), shows that the geographic concentration of the semiconductor industry in the USA increased until the 1980s and then declined. Audretsch and Feldman (1996) compare the location of innovative activity of 210 industries that are in different phases of the industry life cycle. They found out that geographically concentrated companies do exhibit a disproportionately high innovation rate during the growth phase of the industry. Conversely, companies outside clusters are more innovative during later stages. They conclude that "the positive agglomeration effects during the early stages of the industry life cycle are replaced by congestion effects during the later stages of the industry life cycle" (Audretsch and Feldman, 1996: 253). Pouder and St. John (1996) apply an ecological approach to illustrate the movement of the cluster through the life cycle. At the beginning, clusters grow more strongly than the rest of the industry through a creative environment. As it develops, the mental models of the actors become increasingly focused on the (previously) successful trajectory. The former clustering advantages may turn into disadvantages as the clustered companies become locked into a trajectory that once marked their success, but is not able to cope with contemporary development. The decrease in innovative activity in later stages marks the decline of clusters and their development in these phases is worse than that of non-clustered companies (Pouder and St. John, 1996).

What all these insights have in common is that the concentration of the industry increases in younger stages and the prevalence of clusters decreases in mature stages. In doing so, they point to a cluster life cycle that differs from the industry life cycle. Rigby and Essletzbichler (2006) give some indications that can characterize this difference. They compare the heterogeneity of production technologies for three industries on the state level in the USA in different phases of the life cycle: meat packaging, sewing machines, and surgical instruments. They found no evidence of a convergence of production technologies on the national level and heterogeneity persists in all three industries over time. However, when analyzing differences in production technologies on the state level, "plants located in the same state tend to employ production techniques that are relatively similar to one another compared to plants located in different states" (Rigby and Essletzbichler, 2006: 66). Therefore, companies converge within states and remain heterogeneous between states. Surgical instruments, as the youngest industry, are an exception. Rigby and Essletzbichler (2006) argue that it takes time for regional distinctions to emerge. Thus, heterogeneity persists in all industries in different phases of the industry life cycle, but the distribution of heterogeneity changes within states at younger stages and between states in mature stages.

These general results are also confirmed by case studies. As the properties of functioning clusters are well known (Malmberg and Maskell, 2002; Enright, 2003), we concentrate on the presentation of results on the emergence and decline of clusters. Few studies exist on the emergence of clusters,[1] mainly because an emerging cluster is hard to detect and can sometimes only be described in hindsight, as in Bresnahan *et al.* (2001). During its emergence, the cluster is not actually a cluster. But in this phase, the basis for the cluster and subsequent growth processes

[1]The recent book by Braunerhjelm and Feldman (2006) is a notable exception.

is generated. Krugman (1991) claims that "historical accidents" are responsible for cluster emergence. Arthur (1994) describes the emergence of clusters as a stochastic process of start-up and spin-off processes. Spin offs form randomly in different regions. The cluster establishes itself in the region in which the number of companies first exceeds a certain threshold and generates increasing returns. Klepper (2001a) offers a more firm-centered perspective. Successful companies pass their routines on to their spin offs, which then also grow at a disproportionately high rate. Accordingly, the cluster develops in the location where companies with better routines have been formed.

While these approaches consider the clustering process and thus the location of the cluster to be accidental, others also take the characteristics of the local environment into account. Boschma and Wenting (2007) refine Arthur's model. They claim that locations with older but related industries have a higher likelihood of forming a cluster (see also Klepper, 2007 on TV in USA). This is concordant with Jacobs (1969: ch. 4), who states that for the generation of new companies and industries "new work is added to older work". Other studies point out the importance of a strong scientific base as a prerequisite for the emergence of a cluster (Zucker *et al.*, 1998). These approaches have in common that the emergence of clusters requires regional particularities, but in which of these regions the cluster eventually emerges is again completely random. This strong focus on coincidence is countered by Martin and Sunley (2006). They advocate a closer view on the causality of these "coincidences". The development of new regional paths is not fully haphazard or accidental, but can be the result of strategic purpose. This purpose only seems to be accidental, as the logic of the strategic decision, i.e. the "coincidence", differs from the logic of old pathways. Nevertheless, despite their differences, all these contributions share the view that of the many locations with an emerging cluster, only some will bear a functioning and growing cluster.

One of the main factors that turn emerging clusters into a growing one are pervasive spin-off processes (Feldman *et al.*, 2005; Klepper 2007). Although a gradual transition from an emerging cluster to a functioning and growing cluster is possible, during the growth phase of markets, in which the existing companies are well positioned, the crucial push often stems from isolated events and sudden changes in exogenous factors like alterations in the leading companies' organizational structures (Longhi, 1999; Feldman, 2001, Fleming and Frenken 2007) or a change in the technological framework (Dalum *et al.*, 2005). Other insights are given by studies with a micro-oriented perspective. The formation of clusters does not depend on existing companies and an appropriate environment alone, but more on the relations between the companies. Orsenigo (2001) describes how biotechnology companies in Lombardy in fact tended to cluster, but that this (emerging) cluster failed to reach critical mass because, among other reasons, the heterogeneity of its companies was too great. Menzel (2005) emphasizes the importance of developing focal points in clusters. These focal points are based on spin-offs that have the same origin. Due to inheritance of similar routines from their common source, these companies are technologically close and are also connected through various social networks.

Declining clusters, in contrast, have quite different properties. Grabher (1993) describes the coal and steel district in the Ruhr Area as one example of a region that "became 'locked-in' by the very socioeconomic conditions that once made these regions 'stand out against the rest' [... and ...] fell into the trap of 'rigid specialisation'" (Grabher, 1993: 256). He argues that the lock-in comprises several dimensions like technology, network structure, and policy. The decline of clusters like the textile industry in Manchester and the automobile industry in Detroit was also caused by their former success that led to mono-structured "company towns" with too little heterogeneity and diversity to generate new ideas (Jacobs, 1969). These examples indicate that clusters decline when they lose their ability to adjust to a changing environment and that this ability depends on the diversity of knowledge in the cluster.

However, very few clusters follow a rigid life cycle from emergence to growth and decline. Clusters display long-term growth if they are able to maintain their diversity. Grabher and Stark (1997) argue that the success of the industrial districts of the Third Italy is generated by their variety of organizational forms that enables the districts to easily adjust to external changes. In Silicon Valley, a threatening decline of the semiconductor industry due to increasing rigidity of the large mass-producing chipmakers resulted in a new wave of spin-offs producing small batches of highly specialized chips, thereby increasing the diversity of the cluster (Saxenian, 1990). There are also examples of clusters renewing themselves and entering new growth phases. Klepper (2007) describes how the clusters of radio producers in the USA shifted to television. Grabher (1993) illustrates how a new technology path in environmental technologies is emerging in the old coal and steel cluster in the Ruhr Area. The companies of the steel and coal complex acquired the necessary core competencies correcting environmental damage that was caused by their industry. Tappi (2005) illustrates how the accordion cluster in Marche/Italy moved into electronics. The new knowledge was brought into the cluster by expatriates from Marche who had lived in the USA. Clusters are therefore able to enter new life cycles in other industries and leave a maturing industry.

Figure 1 summarizes the stylized facts. It describes the different developments of clustered and non-clustered companies throughout the industry life cycle. When the industry first emerges, the companies are scattered. The few agglomerations of companies in the new industry do not show a better performance than the rest of the industry. But during this phase the foundation for the later clusters is laid. The growth of the industry is accompanied by the increasing preeminence of one or more clusters and a better performance of clustered companies that results in an increasing concentration of the industry. As the industry matures, the development of clustered companies falls behind that of non-clustered ones. A dispersion of economic and innovative activity is a result, which nevertheless might find place

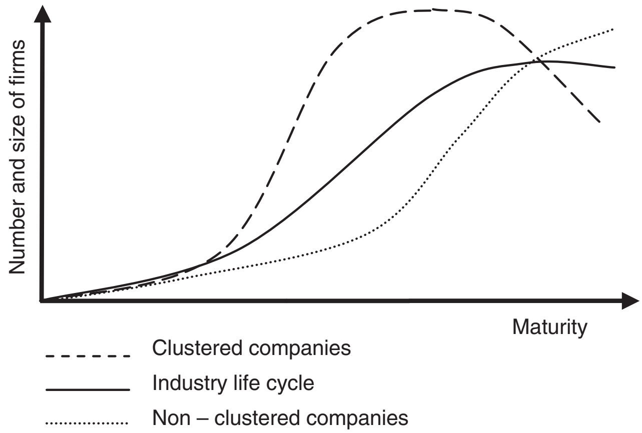

Figure 1 Clustered and non-clustered companies during the industry life cycle.

in the emergence of new clusters. Yet, old clusters may exhibit above average rates of innovation and productivity (Porter, 2003). The figure shows that the cluster dynamics only work positively between two points. The first point is after the cluster's emergence, when there are a sufficient number of companies to reach a critical mass and cluster dynamics start to show an effect. The second point is when the cluster dynamics stop working or have a negative effect on the companies in the cluster, which results in a relative decline. But clusters can also renew themselves. By integrating and applying new technologies and knowledge, they move back to an earlier phase of the cluster life cycle and can enter new growth phases.

3. Elements of clusters

Before we begin with the theoretical details, it is necessary to take one step back and answer the question of what clusters actually are. Although it is not possible to answer this question conclusively, due to the many definitions and forms of clusters (Martin and Sunley, 2003), an approximation of the basic structure of the concept is necessary. There are two reasons for this. The first is analytical. A multitude of qualitative and quantitative studies describe a specific cluster as the result of the interaction of different variables such as regional assets, cluster and industrial dynamics, and socioeconomic conditions. A clarification of the elements and the processes that occur within clusters requires a heuristic that, where possible, separates the factors that are internal to the cluster from those factors that contribute to shaping the cluster, but nevertheless are external and therefore not part of the dynamics of that specific cluster. The second reason is epistemological. The geographical cluster as such is only a theoretical concept that describes a particular constellation of different elements (e.g. companies, organizations, and networks) and processes (e.g. interaction and monitoring). Furthermore, the movement of the cluster through the life cycle is not performed by the cluster, which is only a concept, but is the result of the activities and the evolution of its elements. It is therefore necessary to describe the different elements of clusters to avoid reification.[2]

Porter (1998: 78) defines a cluster as follows: "Clusters are geographic concentrations of interconnected companies and institutions in a particular field". Although this definition is very general, it contains the essential elements of a cluster. First, the cluster consists not only of companies, but also of institutions.[3] These institutions include not only beneficial associations such as cluster organizations, but also research and educational facilities that are the basis for innovation networks and human capital. Companies and institutions are therefore the basic units of a cluster and they cannot be separated from each other, since their respective developments are closely connected (Kenney and Von Burg, 1999; Maskell, 2001).

Secondly, the definition only involves certain companies and institutions. Clusters thus have an outer boundary. Only those companies and institutions belong to the cluster, which are "in a particular field" and exhibit a certain technological proximity or work on a common theme or along a value chain which represents the basis for various exchange processes, synergies and complementarities. Companies and organizations that deal with other themes are outside this "particular field", even if they share the same location. Thus, the cluster is a part of the regional production system that differs from other companies or clusters in the region in its thematic boundary. The thematic boundary can even delimit variations in developmental logic at the same location. Bathelt's (2001) description of the milieu of Boston's emerging biotechnology industry, for example, differs strongly from Saxenian's (1994) picture of the minicomputer industry in the same region. Therefore, the thematic boundary distinguishes the cluster from other parts of the production and innovation system within its spatial range. Apart from the thematic boundary of the cluster, there is also a spatial one, as the companies and their institutional environment are "geographically concentrated". The spatial boundary separates the cluster from its industrial environment, which consists of companies and organizations in the same thematic field which are located elsewhere. Local path dependencies (Martin and Sunley, 2006), transaction costs (Mccann and Sheppard, 2003), small cognitive distances within and large cognitive distances between clusters (Maskell, 2001) create regional characteristics that separate the respective industries geographically (Rigby and Essletzbichler, 2006).

[2]See Lane *et al.* (2006) for the problem of reification with regard to the concept of "absorptive capacity" (Cohen and Levinthal, 1990).

[3]Institutions in the sense Porter defines them refer to organizations (besides firms) and not to social institutions such as social norms or conventions. We follow his definition in this article.

The third and last aspect of Porter's (1998) definition refers to the fact that the companies and institutions within the thematic and spatial boundary are "interconnected". These interconnections refer both to traded and un-traded interdependencies such as market exchange of goods and services, labor market mobility, imitation of behavior, social networks, and face-to-face interaction and cooperation, most of which require a high level of mutual trust and technological proximity as well as horizontal and vertical complementarities of technological activities. These interconnections define which economic sectors the cluster consists of and what geographic extension it has (Porter, 2003). However, relations and exchange processes are not evenly distributed within a cluster (Giuliani, 2007). A high density of interconnections around specific topics forms focal points of activity.[4] Depending on the level of abstraction, a cluster can form around one focal point or consist of several focal points, which in turn may form distinct subclusters. For example, the storage media, software, and semiconductor clusters in Silicon Valley are focal points (or subclusters) of the greater computer cluster. Additionally, interconnections are not only thematically focused, but also spatial. Places with dense activity are often business incubators in younger clusters, with shared facilities of clustered companies, anchor companies, or focal companies in older ones (Feldman, 2003, Fleming and Frenken, 2007). They provide the spatial context in which the actual exchange processes and interactions take place. The necessity for firms to be near the geographical center of the cluster increases with their need to regularly interact face-to-face. This is especially evident for creative industries like media or cultural industries that often concentrate in single city quarters (Scott, 2002). Therefore, the cluster itself consists of a critical mass of companies and institutions around a thematic and spatial focal point. Accordingly, the boundaries of the cluster are marked by a sudden decrease of activity and interconnections at a particular range from the cluster's center.

Figure 2 summarizes the elements of a cluster. The cluster consists of interconnected companies and institutions around a focal point within a spatial and thematic boundary. For a dynamic approach to clusters, the changes these elements are subject to must also be considered; namely, how do the spatial and technological boundaries and the center of the cluster change, how do the interconnections change, and finally, how do the companies and institutions change as the cluster evolves? Apart from the elements of the cluster, Figure 2 reveals both the overarching developmental logic that has to be included in the model and those factors, which are exogenous. The figure includes four types of company and institution: those that belong to the same thematic field, those that share the same location, those that are both in the same thematic field and the same location and hence constitute the cluster, and

[4]Intensity as a measure of delimitation is also applied in other fields, for example in the generation of functional regions where planning or labor market areas are defined by an increase in commuter streams between regions as a proxy for interconnections of economic activities.

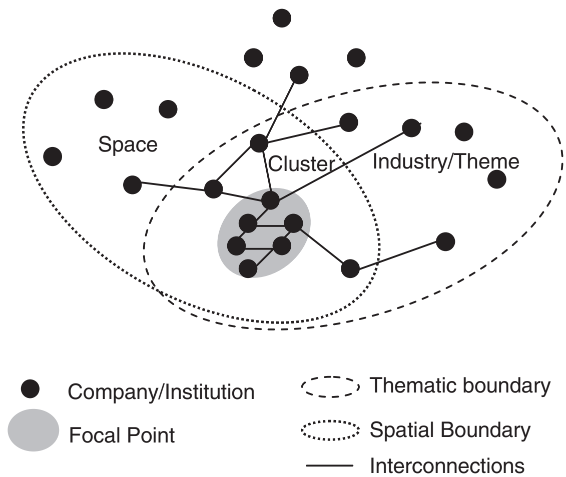

Figure 2 Elements of clusters.

companies and institutions that belong to different industries in different locations. As it is not our intention to discuss all influences on cluster development in depth,[5] we limit ourselves to those factors and processes which are created within the cluster and distinguish clustered from non-clustered companies. We omit dynamics that are created elsewhere and additionally influence the three types of non-clustered companies shown in Figure 2. First, we leave out factors that affect all four categories of company, e.g. the ongoing process of globalization or changes in the socioeconomic paradigm and the contemporary phase of capitalism (Boyer and Durand, 1997; Dicken, 1998) as well as firm specific capabilities and routines (Klepper, 2007). Then we exclude factors that influence all companies in the same specific spatial context such as legislation, size and type of the region, social institutions, and conventions as well as regional assets. A regional culture like that of the Italian industrial districts or of Silicon Valley, for example, influences all companies within the region, whether they are part of a cluster or not. Finally, we omit factors that are limited to a particular industry. Examples of this are possibilities for division of labour (Steinle and Schiele, 2002), market structures such as the oligopolistic market in the pharmaceutical industry or a heterogeneous demand (Malerba *et al.*, 2007) which can prevent the cluster from undergoing a negative lock in.[6] For the same reason, different types of clusters, like those described by Markusen (1996) are

[5]For an overview of the processes that influence industrial evolution, see Malerba (2006). For an overview of regional factors, see Moulaert and Sekia (2003) as well as Brenner and Gildner (2006).

[6]We include demand insofar as it refers to a buyer–supplier relationships within the cluster. Their importance for cluster development is indicated by case studies (Ketelhohn, 2006).

excluded here. Empirical work indicates that the structure and type of the regional cluster is more a result of industrial or market development, and therefore external developments, than of its internal dynamics (Albino *et al.*, 1999). While all these factors have an influence on the emergence, growth and decline of the cluster, we concentrate on the endogenous cluster dynamics that arise from the position of the cluster at the interface between industrial and local dynamics. Describing this interface requires an analysis of its constituting rationales, namely industrial and spatial dynamics and how their synthesis creates the specific cluster dynamics that form the basis for the cluster life cycle model.

4. Industry life cycles and clusters

Several contributions describe industrial dynamics at different stages of development (Vernon, 1966; Markusen, 1985; Klepper, 1997). Like a product, an industry follows cyclical development patterns. Klepper (1997: 148) distinguishes three different stages of an industry life cycle: embryonic, growing and mature:

> [...] In the initial, exploratory or embryonic stage, market volume is low, uncertainty is high, the product design is primitive, and unspecialised machinery is used to manufacture the product. In the second, intermediate or growth stage, output growth is high, the design of the product begins to stabilize, product innovation declines, and the production process becomes more refined as specialized machinery is substituted for labour. Entry slows and a shakeout of producers occurs. Stage three, the mature stage, corresponds to a mature market. Output growth slows, entry declines further, market shares stabilize, innovations are less significant, and management, marketing, and manufacturing techniques become more refined. [...]

In this model, there is an embryonic stage with small output, a subsequent growing stage, and a mature stage with a decline in the number of companies and employees. But neither the age, nor the quantitative development of companies and employees sufficiently describes the development of an industry. Other concepts such as that of the dominant design (Abernathy and Utterback, 1978; Suarez, 2004) and the technological trajectory (Dosi, 1988) point to qualitative factors like competencies and knowledge, which strongly characterize the development of industries and companies. During the emergence of a new technological trajectory, alternative concepts compete with each other. The technological field is quite heterogeneous and there is a large amount of uncertainty concerning the future direction of the trajectory. Over time, certain directions crystallize as the most promising ones, while other possibilities are rejected. By abandoning possibilities that are considered to be inferior, e.g. by shakeout processes and a decrease in the number of companies

(Klepper, 2007), and because existing and newly formed companies orient themselves toward the most promising development approaches, the heterogeneity declines and the technology path becomes increasingly focused. This stabilization and focusing reduces the uncertainty of the further development and is accompanied by a growth of the respective industry (Suarez, 2004). As industries mature, these dynamics decrease. Nevertheless, most mature industries are able to maintain a certain degree of heterogeneity (Rigby and Essletzbichler, 2006; Klepper, 1997). If this degree cannot be maintained, the technological trajectory becomes increasingly narrow and the companies linked to this trajectory can lose their ability for renewal and adaptation. The result is a decline in the corresponding industry. Therefore, there is a quantitative development of the industry described by the number of companies, employees or turnover and a qualitative development, which incorporates the diversity of knowledge and competencies in the industry.

Cluster development resembles the development of the industry life cycle in several ways.[7] As a rule, many approaches describe the stage of the cluster using its age and growth in analogy to the industry life cycle (Enright, 2003; Dalum *et al.*, 2005; Maggioni, 2002). This similarity holds not only for the quantitative development of industries and clusters, but also for their qualitative development.[8] Even if concepts like technological trajectory and dominant design cannot be assigned on a one-to-one basis, they nevertheless indicate a connection between heterogeneity and growth that can also be found in clusters. Tichy (2001) describes this relation using a "cluster paradox". On the one hand, a narrow specialization of the cluster increases the possibility of making use of technological synergies between the companies. On the other hand, this strong similarity of the companies bears the danger of a negative lock-in as it decreases the probability of more radical innovations, which would lead to a wider development path and an increased ability of the cluster to adapt to changing external conditions. The same mechanism applies if the heterogeneity in the cluster is too large. Large diversity facilitates the adjustment of the companies or a change of the development path, but too much heterogeneity can also prevent the exploitation of synergies between companies. Frenken *et al.* (2007) describe with the term "related variety" the importance of an appropriate mean

[7]Theories like the product cycle (Vernon, 1966) or the profit cycle (Markusen, 1985) describe the changing location patterns of industries as dependent on external factors like market structure, investments, or movements from product to process innovation. While our intention is to disentangle industry and cluster logics, these contributions nevertheless point to the similarities and interdependencies between cluster and industry life cycles.

[8]Most cycle theories (Vernon, 1966; Klepper, 1997) propose a change from product innovation to process innovation as the industry matures. A direct analogy would deterministically imply the decline of clusters, as cost reduction becomes more important during later stages of the life cycle. However, the emergence of the industrial districts of the Third Italy in the 1980s, which all produce traditional products, indicates that product innovation can be a crucial source of development during later stages of the industry life cycle as well (Storper, 1985).

between too much heterogeneity on the one side and too much concentration on the other side.

In addition to this, the diversity of clusters must refer to their size. Large clusters such as the computer cluster in Silicon Valley consist of thousands of companies with a great diversity of technologies and knowledge. But the industrial districts of the Third Italy also represent clusters, although these are smaller and of a very specific form. The companies of the industrial district in Carpi, for example, are strongly specialized in their focus on woolens. Nevertheless, both Carpi and Silicon Valley are regarded as positive examples of regional clusters. The size and the technological diversity of clusters must therefore correspond to each other. Because of their size, large clusters can accommodate a greater diversity than small clusters while still generating sufficient synergies between companies. Accordingly, smaller clusters must be strongly focused to be able to utilize synergies.

We describe the ratio of diversity to size using the terms focused and heterogeneous. A small cluster that covers the same technological areas as a large cluster is therefore technologically more heterogeneous than the large cluster. Figure 3 may clarify this relation. The figure shows the companies and institutions of three clusters within a technological space. For reasons of simplification, every company has the same size and capacity to innovate. The three clusters differ in three points. The first is size. Clusters A and B have the same size, while cluster C consists of twice as many companies. The second point is diversity. Diversity is measured by the different knowledge existing in its companies and institutions. Cluster C has the largest diversity due to its larger number of companies. As the clusters A and B consist of the same number of companies and institutions, they bear the same diversity of knowledge. The third difference lies in the heterogeneity that is described by the technological distances between the companies and institutions. The technologically most focused cluster is B. It contains the same number of companies as cluster A, but the companies of cluster A cover wider technological areas and are more heterogeneous than those in cluster B. In contrast, the companies of cluster C bear the

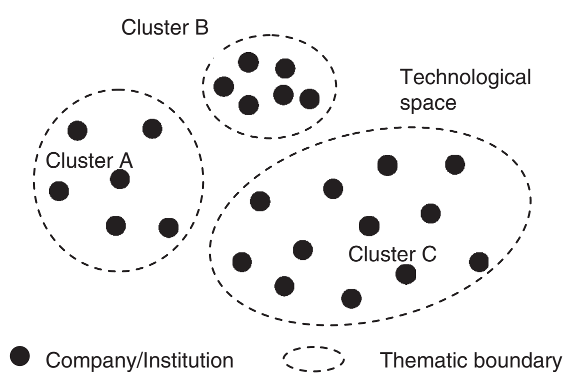

Figure 3 Focused and heterogeneous clusters in a technological space.

largest diversity and cover the widest thematic areas. Nevertheless, cluster C represents the same heterogeneity as cluster A due to its size. Although cluster C covers twice as much technological space as cluster A, it also contains twice as many companies, which compensate for the greater diversity.

The conclusion is that clusters can be distinguished by a quantitative and a qualitative dimension. The quantitative dimension describes the economic development of the cluster in terms of the number of active companies and employees. Because of the possible shift of the cluster into new industries, a description of the cluster according to its development is more appropriate than a description using its age, as old clusters can also grow when they move into new fields. Therefore, we characterize the cluster using the following stages of development: emergence with only a few but growing numbers of mostly small companies, growth with a growing number of employees, and sustainment, when the cluster is able to maintain its employment on a high level in more mature phases. We also add a fourth stage, the declining stage, to account for the fact that a cluster can decline and diminish. In addition to this quantitative account, the qualitative dimension describes the heterogeneity of companies' competencies inherent in the cluster.

The development of the heterogeneity during the stages is illustrated in Figure 4. As the cluster emerges, there are only a few companies and the heterogeneity increases strongly because every new company ventures into new technological areas of the cluster. In the growth phase, the technological path becomes increasingly focused. The heterogeneity decreases until the cluster has matured and a distinct development path has taken shape. However, if the cluster is focused too narrowly, it loses its capacity for renewal and declines. The connections between quantitative and qualitative development of the cluster indicate that its heterogeneity of knowledge is the foundation of its development. The cluster declines if its heterogeneity

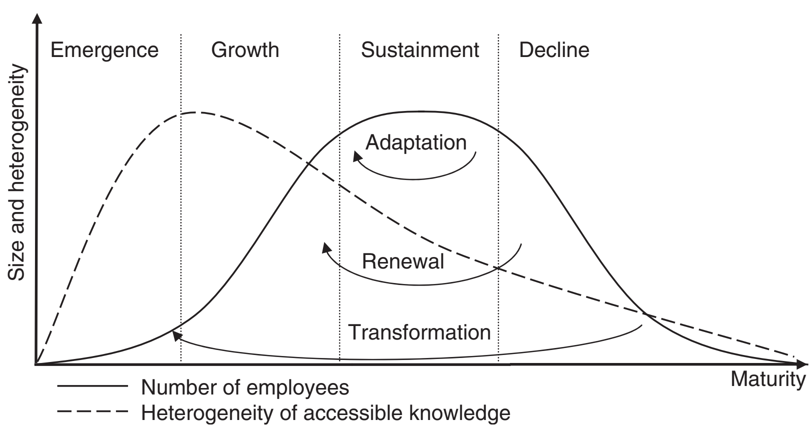

Figure 4 Quantitative and qualitative dimensions of the cluster life cycle.

cannot be sustained. If the heterogeneity increases again, the cluster moves "back" in the cycle and enters a new growth stage. This increase in heterogeneity can be incremental, e.g. the integration of new knowledge from the respective technological trajectory into the cluster. Examples of this are clusters, which manage to maintain their heterogeneity by incrementally adapting to a changing environment. But the increase of heterogeneity can also be of a more radical nature. Clusters can renew themselves by integrating new technologies, like the accordion cluster in Marche/Italy whose companies use electronics in their previously traditional musical instruments (Tappi, 2005). The step back can be larger, when clusters are transformed and move into completely new fields. Such a shift took place in the declining coal and steel complex of the Ruhr Area toward environmental technologies (Grabher, 1993). Additionally, the cluster can increase its heterogeneity by changing its developmental rationale, for example from production to the local organization of global value chains. Therefore, the development of the cluster is not a deterministic move from the left to the right, but a steady oscillation between the left and right sides of the figure.

However, the analogies between industry and cluster life cycle cannot explain why their movement through the cycle differs. Both cycles share heterogeneity as the decisive variable. As a different exploitation of heterogeneity inevitably leads to different developments, it is the utilization of heterogeneity between clustered and non-clustered companies that results in different life cycles (Pouder and St. John, 1996; Audretsch and Feldman, 1996; Rigby and Essletzbichler, 2006).

5. Systemic characteristics of cluster evolution

It has been argued that the driving force of cluster development is the diversity and heterogeneity of knowledge. But this diversity only has an effect if it is put to economic use by the companies. This is precisely what Granovetter (1973) describes with his notion of the "strength of weak ties", i.e. that networks primarily work by making the differences between actors economically exploitable. In her comparative study of the computer industry in Boston and Silicon Valley, Saxenian (1994) demonstrates that (at the time of the study) the advantage of the companies in Silicon Valley lay in the constitution of the local networks that enabled the companies to make use of the diversity of the available skills. Longhi (1999) and Quéré (2007) describe a similar effect in the example of the technology park in Sophia Antipolis. The former "satellite platform district" (Markusen, 1996) was transformed into an operating cluster through spin-off processes. During this process, the knowledge that had existed all the time within the few leading companies, but was not accessible to other companies, was embedded in regional network structures and could subsequently be tapped. Therefore, it is not the mere diversity of knowledge, which is of consequence, but whether or not this knowledge can be put to use.

Cohen and Levinthal (1990) argue that degree to which external knowledge can be exploited depends on the absorptive capacity of the companies. Apart from this company-specific capability, the utilization of knowledge also depends on the strategies, corporate cultures and actions of other companies that make their knowledge accessible. Nonaka and Toyama (2005: 420), for example, see the company as "a dynamic entity which actively interacts with its environment and reshapes the environment, and even itself, through the process of knowledge creation". These two reciprocal effects, i.e. that companies affect and are affected by their environment, are especially applicable to clusters. Companies in a cluster are connected to other companies and institutions of the cluster through various exchange relations and mutual interdependencies (Porter, 1998; Malmberg and Maskell, 2002). As a result, the individual companies and their innovative capabilities are affected by the actions and behavior of other companies and actors of the cluster and, in turn, each individual company and its actions affect the possibilities of other companies to make use of the knowledge. In doing so, even single actions have an influence on the whole cluster system. These effects, which arise through the mutual influence of the different elements of the cluster, are called "systemic effects" below. The systemic effects account for the fact that cluster dynamics emanate from the interconnection between a cluster's different elements and are focused on a complex system of production and innovation.

Systemic effects do not only exist for the qualitative dimension. While the qualitative systemic effects work through the interconnection of knowledge and capabilities, the quantitative systemic effects exert their influence by making use of the number of companies and employees. The size of the cluster, for example, affects the perception of the cluster by external actors. Larger clusters are more likely to be noticed, which can lead to better political support. In contrast, smaller clusters are easily ignored. There are many examples in which political support for the cluster was absent until the cluster reached a specific size. Prevezer (1998) describes how an incubator was set up for the biotechnology cluster in North Carolina after the cluster already had over 100 companies. For a cluster in Washington, DC, Feldman (2001) demonstrates that its development was not initiated by venture capital or other means of support that gained importance later in its development. In addition to this, the regional universities have only offered cluster-related teaching curricula after the emergence of the cluster in the region. Longhi (1999) gives a similar example for Sofia Antipolis. A university with cluster-related educational programs was only set up there after the first growth phase of the cluster. These cases show the importance of perception by external agents and also that this perception depends on the size of the cluster.

Apart from the effects that work externally and contribute to pulling outside actors into the cluster, the systemic quantitative dimension also has an internal affect that becomes apparent in the ability of the companies in the cluster to take joint and collective action to draw attention to their concerns. The larger the number

Table 1 Dimensions of clusters

	Quantitative	Qualitative
Direct	Size (number of organizations, actors, and employees)	Diversity (knowledge, competencies, and organizational forms)
Systemic	Utilization of the size (perception of the cluster, capacity for collective action)	Utilization of the diversity (exploitation of synergies, networks and value chains)

of companies and employees in a cluster, the larger the collective influence these companies have and the more likely it is that they succeed in satisfying their specific needs. An example is given by Saxenian (1994): in the 1940s, a large portion of the national defense spending of the USA went to companies on the East Coast. This relationship deteriorated further when West Coast companies lost some orders. In 1943, an association was formed by Californian companies of the electrical industry "to promote their industry, particularly by lobbying for a share of defence contracts that were going to Eastern companies" (ibid.: 21). This foundation was an expression of the internal capability of taking collective action toward common goals.

The dimensions of a cluster are summarized in Table 1. The quantitative dimension consists of the number of companies and employees, while the qualitative dimension tegrates their specific knowledge and skills into the model. Both dimensions do not only have a direct effect but also influence other parts of the cluster. This aspect is taken into account by the systemic dimension. The quantitative systemic dimension describes the perception of the cluster by external agents as well as the ability of the companies and institutions to take collective action. In addition to this, the qualitative systemic dimension describes the effect that tapping the competencies of the individual companies and institutions has on the whole cluster.

The systemic dimensions also mark the boundaries of the cluster. As the cluster is the result of the interplay between its different elements, the mere existence of companies and organizations as described by the direct quantitative and qualitative dimension is necessary but not sufficient to depict a cluster's boundaries. Maskell (2001: 936) defines the boundaries of the cluster as the "*fit* between the economic activities carried out by the related companies of the cluster on the one hand and the particular institutional endowment developed over time to assist these activities on the other". This *fit*, in the form of the interconnections between the companies and their institutional environment, define both the spatial and thematic boundary of the cluster. The spatial boundary is formed by the quantitative systemic dimension, as this dimension is derived from the location of companies and their capability to build cluster-specific institutions. Actually, the finding of formal institutions is the manifestation of the cluster's capability to turn its size to an advantage. Strong and growing clusters are able to integrate previously unconnected economic activities in distant or adjacent areas into the cluster and, in doing so, enlarge its spatial boundaries. An example of this is Silicon Valley, which originated in the Stanford Industrial Park, grew to encompass several counties, and went on to integrate distant places like Hsinchu/Taiwan (Saxenian and Hsu, 2001) into its developmental logic, while maintaining its geographical center in the Silicon Valley. In addition to the quantitative systemic dimension that describes the spatial boundary of the cluster, the qualitative systemic dimension depicts its thematic boundary. The boundary narrows when companies orient themselves toward each other or actors at the thematic edge leave the cluster. The thematic boundary widens when other actors, e.g. companies or research institutes, are integrated into the cluster. Examples of this are companies that diversify into the cluster or research facilities in related fields that orient themselves toward the topics of the cluster and enlarge the boundaries of the cluster by doing so.

6. Localized learning and cluster dynamics

The previous sections presented the different dimensions of the interconnections between actors in clusters and pointed out that the cause for the different development between clustered and non-clustered companies as well as the movement of the cluster through the different stages lies in the different utilization of heterogeneity between clustered and non-clustered firms. This section illustrates how localization of interaction, or "localized learning" as Malmberg and Maskell (2006) put it; affects the utilization of heterogeneity and, in turn, changes the heterogeneity of the cluster.

Localized learning depends on two factors. The first describes local capabilities that are "some forms of knowledge creation and exchange that are still very much rooted in the cultural, institutional, and social structures of particular places" (ibid.: 3). The second factor involves the influence of spatial proximity on interaction. Since local capabilities refer to local characteristics that apply to all companies in the respective region and hence belong to the catalogue of factors that are outside our model (see Section 2), we concentrate on the second feature of localized learning.

Knowledge exchange in spatial proximity can take place in several ways. One way is direct interaction. Companies can collaborate and in doing so can learn from each other. But direct interaction is not the only way of acquiring external knowledge. Companies monitor each other, observe what their competitors are doing, assess the behavior and results of their competitors' activities, and combine their ideas with their own knowledge. Another way of localized knowledge exchange between companies takes place through social contacts of their employees to employees of other companies or through labor mobility. These different ways of knowledge exchange cause an often diffuse but steady and pervasive stream of information within a region

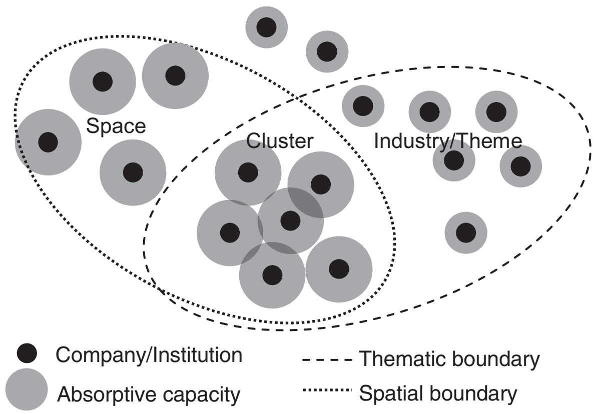

Figure 5 Relative absorptive capacity of co-located and distant firms.

or a cluster, which is described with the word "buzz" by Storper and Venables (2004). Grabher (2002) calls the same effect "noise".

Because of the manifold ways of interactions, the peculiarity of localized learning does not lie in transaction costs or in the efficiency of knowledge transfer. It is rather the diversity of knowledge sources actors that are perceived and accessible and the type of knowledge that is transferred between actors when they are in the same location (Malmberg and Maskell 2006). As a result, companies can learn with and from more heterogeneous actors at a greater technological distance in the same location than with other actors. Empirical results support this argument. Using the example of the biotechnology industry in the USA, Phene *et al.* (2006) determined that technological distances are more easily bridged between companies in the USA than with partners abroad. Zeller (2004) describes hot the big Swiss pharmaceutical companies established research facilities in knowledge centers worldwide in order to tap into the embedded regional knowledge. In the same vein, Owen-Smith and Powell (2004) show that networks connect more easily in spatial proximity. This counts especially for loose networks, where cognitive distance between the nodes is large and thus bear more possibilities for radical innovation (Obstfeld, 2005). Therefore, the absorptive capacity of firms, with what Cohen and Levinthal (1990: 128) describes the "ability of a firm to recognize the value of new, external information, assimilate it, and apply it to commercial ends", depends on the spatial context where at least the recognition and assimilation of external information takes place. Figure 5 illustrates the larger relative absorptive capacity (Lane and Lubatkin, 1998), which accounts for all companies in the same location, whether they belong to a cluster or not.[9]

[9]For reasons of simplification, and to illustrate the main argument, we omit different sizes of companies and different absorptive capacities.

The differences in learning also affect the heterogeneity of firms. During learning, firms do not remain technologically static. When they combine their existing knowledge and resources with new knowledge from external sources, they adjust their knowledge base to the source of the new knowledge and technologically move into its direction (Stuart and Podolny, 1996, Denzau and North, 1994). While this movement through the technological space takes place with all of their collaboration partners independently of their location, it takes place over larger technological distances between firms in geographical proximity, due to their larger relative absorptive capacity (Menzel, 2008).

Due to the extent of technological movement, especially localized learning changes the heterogeneity of the cluster. When companies learn within the cluster, they move in the direction of other companies within the cluster, especially toward successful companies and the cluster's dominant design (Abernathy and Utterback, 1978, Malmberg and Maskell, 2002; Maskell and Malmberg, 2007). Their common focus further facilitates learning processes that again are intensified by cluster-specific institutions (Maskell, 2001). Everything else being equal, the technological distances between companies, and thereby the heterogeneity of the cluster, decrease accordingly. An increase of heterogeneity takes place by learning outside the cluster. Due to the bridged technological distances, this increase is stronger when it takes place locally. Therefore, the larger relative absorptive capacity of co-localized firms in connection with the amount and diversity of knowledge available within this absorptive reach, both within and outside the cluster, distinguishes the development of clustered and non-clustered firms.

7. Stages of the cluster life cycle

The previous article presented the foundations of our cluster life-cycle model. The dimensions of the cluster are the quantitative, qualitative, direct, and systemic dimensions. The dynamics of the cluster are described by the heterogeneity of knowledge, which is responsible for growth, the larger relative absorptive capacity between clustered firms that leads to a better exploitation of heterogeneous knowledge results in a convergence of companies' activities due to mutual learning processes. This article applies these findings to the different stages of the cluster life cycle. We concentrate on those processes that take place within the boundaries of the clusters as defined in Figure 2. It can be difficult to assign a cluster to a specific stage if it is in transition. A cluster consists of many diverse protagonists that develop differently. This diversity means that the cluster does not develop evenly and as a whole. Parts of the cluster can stay at an earlier stage while others advance along the trajectory. Focal points of activity, where most synergies exist between the actors, are supposed to move more quickly through the cycle while the actors at the edge of the cluster lag behind. Therefore, the assignment of the cluster to one of the phases depends on the development of the focal points of the cluster, even though they may change as the cluster moves through the cycle.

7.1. Emerging clusters

It is difficult to precisely define the phase in which a cluster first arises. The main reason for this is that the emerging cluster is not actually a cluster. One example is Silicon Valley in the late 1950s. Emerging clusters only contain a few, often quite small companies with few employees that are scattered over wide areas technologically. Nevertheless, these enterprises can already represent the future technological orientation of the cluster, just as Fairchild Semiconductor and the "Fairchildren" as spin-offs did in the case of Silicon Valley. But their heterogeneity hampers exchange and limits possibilities for local networks, and customer–supplier relations are scarce. Synergies in this phase mainly exist between spin-offs and their parent organization, often a research facility or a university (Shohet, 1998). As a spatial concentration of companies linked by a common topic hardly exists and corresponding economic activity in regions with an emerging cluster can barely be distinguished from regions without a cluster, it is highly probable that an emerging cluster will not be recognized at all. But emerging clusters differ from locations with normal economic activity in two points. The first is grounded in the companies themselves. One or more companies exist which offer a lasting vision for a new local technology path. The second condition lies in the local environment. Certain conditions are given, for example a strong scientific base or political support, which give the emerging cluster the potential to reach a critical mass.

The emerging phase ends in one of two different ways. The emerging cluster either becomes a growing cluster when it is able to reach a critical mass and the growth rate of the companies exceeds the growth rate of non-clustered companies. The crucial factor for this is to first create synergies around a focal point. One possibility is by increasing the number of companies, especially through spin-offs that are regarded as responsible for the first growth of a cluster (Arthur, 1994; Feldman *et al.*, 2005; Klepper, 2007). However, the number of spin-offs is not strictly a function of the size or number of existing companies. The rate of start-ups is also dependent on the quantitative systemic dimension. Social networks, an entrepreneurial environment in the region or organizations that support start-ups influence how many potential spin-offs are actually generated. Another factor that comes to bear on the generation of focal points lies in the qualitative systemic dimension that influences the type of spin-offs. Spin-offs contribute to the generation of synergies in the cluster when they concentrate around thematic focal points (Menzel, 2005).

In addition to this, existing companies are not static but move through the technological space. Due to learning processes, companies reduce their technological distances. This convergence makes further collaboration possible, along with a division of labor and a specialization of the companies and the accompanying opportunities for increasing returns. The part of the cluster where these effects happen at first marks the growth pole of the emerging cluster. Furthermore, the places where these effects come into play will show higher growth rates of companies and have a high potential to form a functioning cluster. The second possibility for ending the stage of emergence is when the emerging cluster loses its potential to become a functioning one, i.e. when its ability to form a focal point vanishes. Two reasons are decisive for this. One would be a thinning out of the already dispersed companies. The companies develop in different technological directions, thereby expanding the technological distances between them. The second reason would be a reduction of mass, for example through bankruptcy or relocation of companies. These "lost" companies leave gaps in the competence structure of the emerging cluster that further limit possibilities for interaction. In the end, the formerly emerging cluster may completely disappear.

7.2. Growing clusters

A strong increase in employment as a result of the strong growth of existing companies and a high number of start-ups characterize a growing cluster. Examples of this stage are, among others, the semiconductor industry in Silicon Valley in the 1970s (Saxenian, 1994) or contemporary biotechnology in Boston (Bathelt, 2001). Unlike the emerging cluster, the boundaries are now definable. Both the existing companies and the start-ups orient themselves toward the growth centers of the cluster. A shakeout of companies at the edge of the cluster additionally decreases the heterogeneity. This convergence further narrows the boundaries of the cluster, the cluster becomes more focused and a "dominant cluster design" forms. The growing density of companies and institutions within the boundaries of the cluster increasingly creates possibilities for innovation networks or customer–supplier relations and forms a specialized labor market. The continuous emergence of new potential networking partners prevents an isolation of individual networks and generates an environment, which has a positive effect on existing companies and start-ups. Improving supportive infrastructure and cluster organizations are established to fulfill and lobby for the fulfillment of the cluster's needs.

The growing stage ends when the growth of a cluster adjusts to the industry average and the cluster arrives at the sustaining stage (Pouder and St. John, 1996), albeit at higher productivity. The main reason for this adjustment is the exhaustion of the diversity by ongoing shakeouts of companies and a more focused orientation of the cluster. Detroit is one example of this. After the growth stage with extensive founding of start-ups and a subsequent shakeout, the cluster achieved its final form with Chrysler, Ford, and General Motors as leading companies (Klepper, 2001b).

7.3. *Sustaining clusters*

The sustaining cluster describes a state of equilibrium. Examples of this are the industrial districts of Baden–Württemberg or the Third Italy (Becattini, 2002; Cooke and Morgan, 1993). A sustaining cluster shows neither a high growth compared to the respective industry nor a remarkable decrease in the number of companies or employees. Fluctuations are more of a cyclical than of a structural nature. The various competencies of the companies are made accessible by dense and established networks. The connections of the companies of the cluster to outside companies and institutions bring new knowledge into the cluster and keep the networks open (Albino *et al.*, 1999; Bathelt *et al.*, 2004). The thematic boundaries of the cluster now move incrementally as new technologies are integrated into the cluster. During its development, the cluster has shaped its regional environment, whose development may even be equated with that of the cluster if it is very dominant (Bresnahan *et al.*, 2001).

There are two ways in which the sustaining stage can end. The first follows the cluster cycle. A decreasing diversity in an exhausted trajectory causes a decline. The second possibility is a step back in the cluster life cycle through the generation of new heterogeneity and an accompanying shift of the thematic boundary, which results in a new growth phase. However, the altering and renewal of the development path often takes the shape of a substantial crisis and therefore comes after a stage of decline (Martin and Sunley, 2006; Meyer-Stamer, 1998).

7.4. *Declining clusters*

A declining cluster is defined by a decrease in the number of companies and especially of employees due to failures, mergers, and rationalizations. Examples of declining clusters are old industrialized regions like the Ruhr Area (Grabher, 1993). Start-ups are rare during this phase. A region with a shrinking cluster is marked by a strong cluster-oriented bias of economic activities. This bias is caused by a specific knowledge base, highly qualified and specialized employees and companies strongly focused on specific markets and technologies. The competencies of such a cluster are contained in only a few companies. Despite the decline, competitive pressure can lead to high innovation rates (e.g. Grabher, 1993; Graham, 1956). These innovations, however, arise within the existing and exhausted technology path and the cluster is negatively locked into its previously successful development path. The reason for a lock-in lies not only in the exhausted regional trajectory, but also in the long-existing, closed, and homogeneous networks, which are unable to renew the cluster by integrating new and often external knowledge. A declining cluster has therefore lost the ability to sustain its diversity, its ability to adjust to changing conditions as well as its potential for an independent renewal. Nevertheless, due to the strong networks, even a declining cluster can be capable of effective collective action, such as lobbying for government support.

There are three possibilities for the declining stage of the cluster to end. The first, again, follows the cluster cycle and the cluster simply diminishes. The other two possibilities break new ground due to an increase in heterogeneity. One of them is a renewal of the existing development path through the implementation of new, yet related technologies, often from other locations. The other possibility of ending the declining stage is a transition toward completely different fields. This implies the integration of completely new actors into the cluster. Due to the intense learning processes necessary for this change, it is probable that the respective actors that bear the new knowledge are located in the same region, but in completely different fields.

Figure 6 summarizes the stages of the cluster life cycle. During the emergence of the cluster, the larger absorptive capacity between the clustered companies enables them to exploit technological distances earlier and make use of more synergies compared to the non-clustered companies. In doing so, they contribute to creating the actual cluster. During the growth phase of the cluster, the companies adjust to each other further and generate optimal technological distances (Wuyts *et al.*, 2005). In doing so, the thematic boundaries of the cluster narrow, as illustrated by the arrows. The implicit tendency to adjust to each other would eventually lead to a sub-optimally small technological distance between the companies (Maskell and Malmberg, 2007). An optimal technological distance can only be sustained by start-ups that deviate from the existing knowledge base (Almeida and Kogut, 1997), by steady implementation of external knowledge that is in line with the trajectory of the cluster or by widening the thematic boundaries of the cluster and integrating new technologies and capabilities into the cluster. If the clustered companies cannot maintain the balance between internal convergence and divergence by what may often be external knowledge, the position in a cluster may negatively affect the companies compared to non-clustered companies. This balance effect through knowledge created elsewhere is not necessary for non-clustered companies as they always rely on external sources of knowledge.

8. Conclusion and discussion

The article introduces a life-cycle model for clusters. We began with the observation that clusters move through different developmental stages and that these developmental stages differ from those of the development of the respective industry. The development of the cluster through the different stages is not only quantitatively described by a growth and decline in the number of companies and employees, but also qualitatively by the diversity and heterogeneity of knowledge. Additionally, there is a systemic dimension that accounts for the fact that companies and organizations in a cluster are part of a complex production and innovation system and that, through various interconnections, both influence and are influenced by other

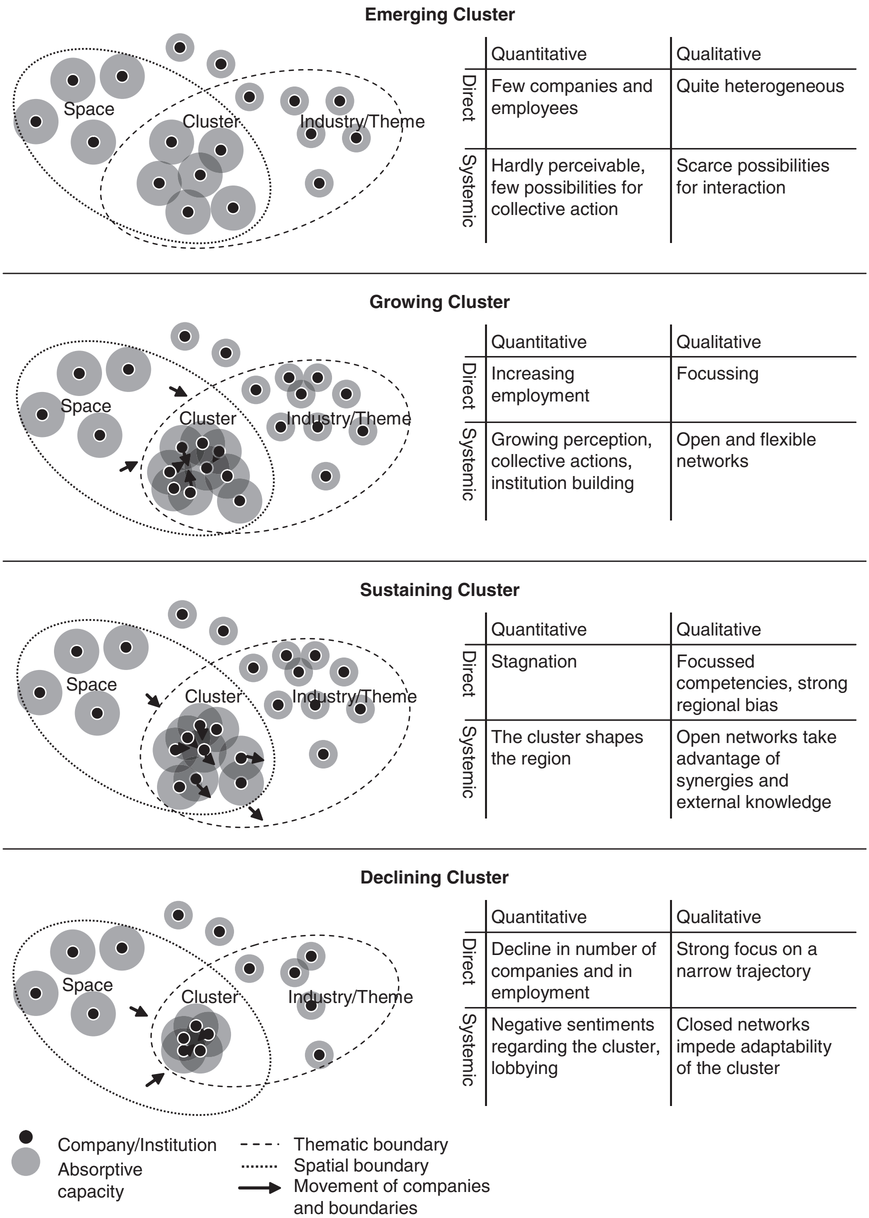

Figure 6 The stages of the cluster life cycle.

companies and organizations. The quantitative systemic dimension illustrates the perception of the cluster by external actors and the capability of the companies to take collective action. The qualitative systemic dimension describes the innovation and learning processes of companies.

We strongly separate between the internal cluster dynamics and exogenous forces to elaborate the distinctiveness of clustered compared to non-clustered firms. As a result, our model deviated from other approaches in three points. The first is the perspective on clusters. While other approaches analyze clusters solely by the absolute or relative number of firms or employees like in Maggioni (2002), Klepper (2001b, 2007), Brenner and Gildner (2007), Prevezer (1998), our model stresses the importance of heterogeneity between the organizations in the cluster, i.e. two clusters with the same size and even in the same industry would be in different phases of the cluster life cycle, if their heterogeneity and the relative position of firms in the technological space differ.

The second difference refers to the utilization of the cluster's heterogeneity. The larger absorptive capacity enables co-located firms to bridge larger technological distances and access more diverse knowledge compared to distant companies. Malmberg and Maskell (2006) already made a similar point when they argued for a smaller cognitive distance between firms in a cluster than between firms in different clusters. We add that this capacity has different implications for firm development, depending on the stage of the cluster life cycle.

The third and perhaps most important difference to existent contributions is the attempt to open the black box of the firm during cluster development. The knowledge of firms' changes as a result of the specific learning processes that take place within clusters. Geographical proximity not only substitutes for technological distances, but continual interaction also results in a steady movement of the companies toward each other. The convergence of firms not only creates additional synergies, but also narrows the thematic boundaries of the cluster. It is this very learning within clusters and the subsequent change of the clustered firm's competencies, which is the driver through the life cycle and distinguishes the development of clustered firms compared to non-clustered firms.

We applied the model to the following four stages: emerging, growing, sustaining, and declining. The model shows that clusters are only advantageous for companies between two particular points. The first point is after the emergence, when the concentration reaches a critical mass and clustered outperform non-clustered companies. The second point is when the companies have exploited the heterogeneity of the cluster and the development of clustered companies is inferior to that of non-clustered companies. These two points mark the crucial questions regarding cluster development, namely how does the emerging cluster become a growing cluster how can the decline of a cluster be prevented.

In addition to existing approaches that argue that the reach of the critical mass mainly depends on the pure number of firms (Arthur 1994, Bottazzi *et al.* 2007), we suggest two additional aspects. Firstly, the heterogeneity between firms must decline to a certain threshold; and secondly, the heterogeneity is affected by changes of the companies' knowledge base. It decreases, when firms adjust toward each other by mutual learning processes. This adjustment of heterogeneous firms toward a common technology or (dominant) design can create first growth poles. This convergence, connected with the larger relative absorptive capacity, enables clustered firms to utilize external knowledge from more and different organizations than non-clustered firms. The access to different knowledge is especially important, when the industry is emerging and an industry-wide dominant design has not emerged (yet). Therefore, we suggest the following hypothesis:

> *H1: Clusters are established in those regions where the knowledge bases of companies converge around technological focal points in the emergence phase.*

Except for the emerging stage, the cluster can steadily renew itself, if firms are sufficiently different to exploit distances between them. Not later than at the sustaining stage, the maintenance of differences between firms becomes the crucial factor. Differences between firms can be created or maintained by two processes. The first refers to the actual firms of the cluster. Linkages to other knowledge flows outside the cluster can pull the clustered firms apart and create new technological distances. Additionally, heterogeneity can be created by integrating new organizations into the cluster that previously have been outside its spatial or technological limitation. By integrating spatially distant but technologically proximate firms, the cluster widens its spatial boundary and gets access to knowledge embedded in other regions. By fostering contacts to other organizations in spatial proximity that previously have been outside its technological boundary, the cluster widens its thematic boundaries and the level of heterogeneity increases. These assumptions lead to the second hypothesis:

> *H2: Clusters can increase heterogeneity and renew themselves by enlarging their boundaries, either by integrating firms in the same industry, but in other places or by integrating organisations in spatial proximity, but outside the thematical focus of the cluster.*

While the enlarging of the cluster's boundaries increases its heterogeneity and leads to its renewal, the actual growth of the cluster depends on the exploitation of the newly generated heterogeneity. Therefore, there is a time lag between the generation of heterogeneity and the growth that results form its exploitation. Due to this reasoning, we postulate the third hypothesis:

> *H3: Firms in established clusters, whose heterogeneity had increased, subsequently outperform firms in steadily narrowing clusters.*

The article provided a detailed discussion of the internal dynamics of the cluster that distinguish clustered from non-clustered firms. In the elaboration of our model, we followed a quite reductionist procedure. We omitted an enumeration of general processes and factors that influence clustered firms, as most of these factors would also affect firms outside clusters. Nevertheless, the cluster life-cycle model could be a starting point to analyze the particular influences of excluded factors like demand, industry structure, regional milieus, different institutional frameworks, and possibilities for division of labor. Methodological approaches to assess these influences could be simulation models like those applied by Vicente and Suire (2008) to analyze the interdependencies between cluster stages and locational behavior of firms.

Beside a deeper theoretical analysis of the underlying processes and closer consideration of the external environment, much research is also necessary to empirically test the model. The life-cycle model consists of four dimensions that each requires distinct empirical methods and data. Köhler and Otto (2006) for example, tested the direct quantitative dimension of the cluster during different stages with regard to start-ups. Their analysis is based on quantitative data including company survival rates, number of start-ups, and number of failures with an eye on employment growth in different clusters at different stages. An examination of the qualitative dimension would have to focus on degrees of heterogeneity with regard to the technological fields the cluster companies are active in over time. This information could be gleaned from patent data or business fields in which firms are active.

The assessment of the systemic dimensions requires more qualitative measures, as their influence only becomes apparent in in-depth case studies. Yet, they are necessary to analyze the co-development of the cluster's size and technological heterogeneity and its institutional environment, for example how an institutional setting changes the technological configuration of a cluster or how new institutions are a result of the evolution of the cluster. Therefore, further empirical work on the cluster life-cycle model would require both quantitative and qualitative research. The necessity of applying both types of research highlights the complexity of this research field. The task at hand is to combine the different kinds of methodological approaches in a single framework. The cluster life-cycle model could contribute to structuring the analytically and methodologically complexity in a meaningful way.

Acknowledgements

We are grateful to the Economic Geography Group at the University of Berne, Sebastian Henn, Ron Boschma, Koen Frenken, Christian Ostergaard, and Rik Wenting for comments on previous versions of the article. We also would like to thank two anonymous referees for their helpful comments to improve and refine the article.

Addresses for correspondence

Max-Peter Menzel, University of Hamburg, Institute of Geography, Bundesstraße 55, 20146 Hamburg, Germany. e-mail: max-peter.menzel@web.de

Dirk Fornahl, BAW Institut für regionale Wirtschaftsforschung GmbH, Wilhelm-Herbst-Straße 5, D-28359 Bremen, Germany. e-mail: d.fornahl@baw-bremen.de

References

Abernathy, W. J. and J. M. Utterback (1978), 'Patterns of industrial innovation,' *Technology Review*, **80**, 40–47.

Albino, V., A. C. Garavelli and G. Schiuma (1999), 'Knowledge transfer and inter-firm relationships in industrial districts: the role of the leader firm,' *Technovation*, **19**, 53–63.

Almeida, P. and B. Kogut (1997), 'The exploration of technological diversity and geographic localization in innovation: start-up firms in the semiconductor industry,' *Small Business Economics*, **9**, 21–31.

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

Audretsch, D. B. and M. P. Feldman (1996), 'Innovative clusters and the industry life cycle,' *Review of Industrial Organization*, **11**, 253–273.

Baptista, R. (2000), 'Do innovations diffuse faster within geographical clusters?,' *International Journal of Industrial Organization*, **18**, 515–535.

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

Bathelt, H. (2001), 'Regional competence and economic recovery: divergent growth paths in boston's high technology economy,' *Entrepreneurship & Regional Development*, **13**, 287–314.

Bathelt, H., G. A. Malmber and P. Maskell (2004), 'Clusters and knowledge: local buzz, global pipelines and the process of knowledge creation,' *Progress in Human Geography*, **28**, 31–56.

Becattini, G. (2002), 'From Marshall's to the Italian "Industrial Districts". a brief critical reconstruction,' in A. Q. Curzio and M. Fortis (eds), *Complexity and Industrial Clusters: Dynamics and Models in Theory and Practice*. Physica-Verlag: Heidelberg and New York, pp. 82–106.

Bergman, E. M. (2007), 'Cluster life cycles: an emerging synthesis,' in C. Karlsson (ed.), *Handbook of Research on Cluster Theory*. Edward Elgar: Cheltenham, UK, pp. 114–132.

Boschma, R. A. and K. Frenken (2006), 'Why is economic geography not an evolutionary science? Towards an evolutionary economic geography,' *Journal of Economic Geography*, **6**, 273–302.

Boschma, R. and R. Wenting (2007), 'The spatial evolution of the British automobile industry: Does location matter?,' *Industrial and Corporate Change*, **16**, 213–238.

Bottazzi, G., G. Dosi, G. Fagiolo and A. Secchi (2007), 'Modeling industrial evolution in geographical space,' *Journal of Economic Geography*, **7**, 651–672.

Boyer, R. and J.-P. Durand (1997), *After Fordism*. MacMillan Pres Ltd.: London.

Braunerhjelm, P. and M. P. Feldman (2006), *Cluster Genesis: Technology-Based Industrial Development*. Oxford University Press: Oxford and New York.

Brenner, T. and A. Gildner (2006), 'The long-term implications of local industrial clusters,' *European Planning Studies*, **14**, 1315–1328.

Bresnahan, T., A. Gambardella and A. Saxenian (2001), 'Old economy' inputs for 'new economy' outcomes: cluster formation in the new silicon valleys,' *Industrial and Corporate Change*, **10**, 835–860.

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

Cooke, P. and K. Morgan (1993), 'The regional innovation system in Baden-Württemberg,' *International Journal of Technology Management*, **9**, 394–429.

Dalum, B., C. O. R. Pedersen and G. Villumsen (2005), 'Technological life-cycles – lessons from a cluster facing disruption,' *European Urban and Regional Studies*, **12**, 229–246.

Denzau, A. T. and D. C. North (1994), 'Shared mental models – ideologies and institutions,' *Kyklos*, **47**, 3–31.

Dicken, P. (1998), *Global Shift: Transforming the World Economy*. Paul Chapman Publishing: London.

Dosi, G. (1988), 'Sources, procedures and microeconomic effects of innovation,' *Journal of Economic Literature*, **26**, 1120–1171.

Ellison, G. and E. L. Glaeser (1997), 'Geographic concentration in US manufacturing industries: a dartboard approach,' *Journal of Political Economy*, **105**, 889–927.

Enright, M. J. (2003), 'Regional clusters: what we know and what we should know,' in J. Bröcker, D. Dohse and R. Soltwedel (eds), *Innovation Clusters and Interregional Competition*. Springer: Berlin, pp. 99–129.

Feldman, M. (2003), 'The locational dynamics of the us biotech industry: knowledge externalities and the anchor hypothesis,' *Industry and Innovation*, **10**, 311–329.

Feldman, M. P. (2001), 'The entrepreneurial event revisited: firm formation in a regional context,' *Industrial and Corporate Change*, **10**, 861–891.

Feldman, M. P., J. Francis and J. Bercovitz (2005), 'Creating a cluster while building a firm: entrepreneurs and the formation of industrial clusters,' *Regional Studies*, **39**, 129–141.

Fleming, L. and K. Frenken (2007), 'The evolution of inventor networks in the Silicon Valley and Boston regions,' *Advances in Complex Systems*, **10**, 53–71.

Frenken, K., F. Van Oort and T. Verburg (2007), 'Related variety, unrelated variety and regional economic growth,' *Regional Studies*, **41**, 685–697.

Giuliani, E. (2007), 'The selective nature of knowledge networks in clusters: evidence from the wine industry,' *Journal of Economic Geography*, **7**, 139–168.

Grabher, G. (1993), 'The weakness of strong ties. the lock-in of regional development in the Ruhr area,' in G. Grabher (ed.), *The Embedded Firm*. Routledge: London, pp. 255–277.

Grabher, G. (2002), 'The project ecology of advertising: Tasks, talents and teams,' *Regional Studies*, **36**, 245–262.

Grabher, G. and D. Stark (1997), 'Organizing diversity: evolutionary theory, network analysis and postsocialism,' *Regional Studies*, **31**, 533–544.

Graham, G. S. (1956), 'The ascendancy of the sailing ship 1850–85,' *The Economic History Review*, **9**, 74–88.

Granovetter, M. S. (1973), 'Strength of weak ties,' *American Journal of Sociology*, **78**, 1360–1380.

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

Kenney, M. and U. Von Burg (1999), 'Technology, entrepreneurship and path dependence: industrial clustering in Silicon Valley and Route 128,' *Industrial and Corporate Change*, **8**, 67–103.

Ketelhohn, N. W. (2006), 'The role of clusters as sources of dynamic externalities in the US semiconductor industry,' *Journal of Economic Geography*, **6**, 679–699.

Klepper, S. (1997), 'Industry life cycles,' *Industrial and Corporate Change*, **6**, 145–181.

Klepper, S. (2001a), 'Employee startups in high-tech industries,' *Industrial and Corporate Change*, **10**, 639–674.

Klepper, S. (2001b), *The Evolution of the U.S. Automobile Industry and Detroit as its Capital*. Carnegie Mellon University: Mimeo.

Klepper, S. (2007), 'The evolution of geographic structures in new industries,' in K. Frenken (ed.), *Applied Evolutionary Economics and Economic Geography*. Edward Elgar: Cheltenham, UK, pp. 69–92.

Köhler, S. and A. Otto (2006), The role of new firms for the development of clusters in Germany. Institute of Employment Research (Saarbrücken/Germany): Mimeo.

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

Lane, P. J., B. R. Koka and S. Pathak (2006), 'The reification of absorptive capacity: a critical review and rejuvenation of the construct,' *Academy of Management Review*, **31**, 833–863.

Lane, P. J. and M. Lubatkin (1998), 'Relative absorptive capacity and interorganizational learning,' *Strategic Management Journal*, **19**, 461–477.

Longhi, C. (1999), 'Networks, collective learning and technology development in innovative high technology regions: the case of Sophia Antipolis,' *Regional Studies*, **33**, 333–342.

Lorenzen, M. (2005), 'Why do clusters change?,' *European Urban and Regional Studies*, **12**, 203–208.

Maggioni, M. A. (2002), *Clustering Dynamics and the Location of High-Tech-Firms*. Physica-Verlag: Heidelberg and New York.

Malerba, F. (2006), 'Innovation and the evolution of industries,' *Journal of Evolutionary Economics*, **16**, 3–23.

Malerba, F., R. R. Nelson, L. Orsenigo and S. G. Winter (2007), 'Demand, innovation, and the dynamics of market structure: the role of experimental users and diverse preferences,' *Journal of Evolutionary Economics*, **17**, 371–399.

Malmberg, A. and P. Maskell (2002), 'The elusive concept of localization economies: towards a knowledge-based theory of spatial clustering,' *Environment and Planning A*, **34**, 449.

Malmberg, A. and P. Maskell (2006), 'Localized learning revisited,' *Growth and Change*, **37**, 1–18.

Markusen, A. (1985), *Profit Cycles, Oligopoly, and Regional Development*. MIT Press: Boston, MA.

Markusen, A. (1996), 'Sticky places in slippery space: a typology of industrial districts,' *Economic Geography*, **72**, 293–313.

Martin, R. and P. Sunley (2003), 'Deconstructing clusters: chaotic concept or policy panacea?,' *Journal of Economic Geography*, **3**, 5–35.

Martin, R. and P. Sunley (2006), 'Path dependence and regional economic evolution,' *Journal of Economic Geography*, **6**, 395–437.

Maskell, P. (2001), 'Towards a knowledge-based theory of the geographical cluster,' *Industrial and Corporate Change*, **10**, 921–943.

Maskell, P. and A. Malmberg (2007), 'Myopia, knowledge development and cluster evolution,' *Journal of Economic Geography*, **7**, 603–618.

Mccann, P. and S. Sheppard (2003), 'The rise, fall and rise again of industrial location theory,' *Regional Studies*, **37**, 649–663.

Menzel, M.-P. (2005), 'Networks and technologies in an emerging cluster: the case of bioinstruments in Jena,' in C. Karlsson, B. Johansson and R. R. Stough (eds), *Industrial Clusters and Inter-firm Networks*. Edward Elgar Publishing: Cheltenham, UK and Northampton, MA, pp. 413–449.

Menzel, M.-P. (2008), Dynamic proximities: changing relations by creating and bridging distances,' *Papers in Evolutionary Economic Geography* 0814. Utrecht University.

Meyer-Stamer, J. (1998), 'Path Dependence in regional development: persistence and change in three industrial clusters in Santa Catarina, Brazil,' *World Development*, **26**, 1495–1511.

Moulaert, F. and F. Sekia (2003), 'Territorial innovation models: a critical survey,' *Regional Studies*, **37**, 289–302.

Nonaka, I. and R. Toyama (2005), 'The theory of the knowledge-creating firm: subjectivity, objectivity and synthesis,' *Industrial and Corporate Change*, **14**, 419–436.

Obstfeld, D. (2005), 'Social networks, the Tertius Lungens and orientation involvement in innovation,' *Administrative Science Quarterly*, **50**, 100–130.

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

Orsenigo, L. (2001), 'The (failed) development of a biotechnology cluster: the case of Lombardy,' *Small Business Economics*, **17**, 77–92.

Phene, A., K. Fladmoe-Lindquist and L. Marsh (2006), 'Breakthrough innovations in the US biotechnology industry: the effects of technological space and geographic origin,' *Strategic Management Journal*, **27**, 369–388.

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

Porter, M. E. (1998), 'Clusters and the new economics of competition,' *Harvard Business Review*, **76**, 77–90.

Porter, M. E. (2003), 'The economic performance of regions,' *Regional Studies*, **37**, 549–578.

Pouder, R. and C. H. ST. John (1996), 'Hot spots and blind spots: geographical clusters of firms and innovation,' *Academy of Management Review*, **21**, 1192–1225.

Prevezer, M. (1998), 'Clustering in biotechnology in the USA,' in G. M. P. Swann, M. Prevezer and D. Stout (eds), *The Dynamics of Industrial Clustering – International Comparisons in Computing and Biotechnology*. Oxford University Press: New York, pp. 124–193.

Quéré, M. (2007), 'Sophia-Antipolis as a 'reverse' science park: from exogenous to endogenous development,' in K. Frenken (ed.), *Applied Evolutionary Economics and Economic Geography*. Edward Elgar: Cheltenham, UK and Northampton, MA, pp. 48–68.

Rigby, D. L. and J. Essletzbichler (2006), 'Technological variety, technological change and a geography of production techniques,' *Journal of Economic Geography*, **6**, 45–70.

Saxenian, A. (1990), 'Regional networks and the resurgence of Silicon Valley,' *California Management Review*, **33**, 89–112.

Saxenian, A. (1994), *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press: Cambridge, MA.

Saxenian, A. and J.-Y Hsu (2001), 'The Silicon Valley–Hsinchu connection: technical communities and industrial upgrading,' *Industrial and Corporate Change*, **10**, 893–920.

Scott, A. J. (2002), 'A new map of Hollywood: the production and distribution of American motion pictures,' *Regional Studies*, **36**, 957–975.

Shohet, S. (1998), 'Clustering and UK biotechnology,' in G. M. P. Swann, M. Prevezer and D. Stout (eds), *The Dynamics of Industrial Clustering – International Comparisons in Computing and Biotechnology*. Oxford University Press: New York, pp. 194–224.

Steinle, C. and H. Schiele (2002), 'When do industries cluster? A proposal on how to asses an industrie's propensity to concentrate at a single region or nation,' *Research Policy*, **31**, 849–858.

Sternberg, R. and T. Litzenberger (2004), 'Regional clusters in Germany – their geography and their relevance for entrepreneurial activities,' *European Planning Studies*, **12**, 767–791.

Storper, M. (1985), 'Oligopoly and the product cycle – essentialism in economic-geography,' *Economic Geography*, **61**, 260–282.

Storper, M. and A. J. Venables (2004), 'Buzz: face-to-face contact and the urban economy,' *Journal of Economic Geography*, **4**, 351–370.

Stuart, T. E. and J. M. Podolny (1996), 'Local search and the evolution of technological capabilities,' *Strategic Management Journal*, **17**, 21–38.

Suarez, F. F. (2004), 'Battles for technological dominance: an integrative framework,' *Research Policy*, **33**, 271–286.

Suire, R. and J. Vicente (2009), 'Why do some places succeed while others decline? A social interaction model of cluster viability,' *Journal of Economic Geography*, **9**, 381–404.

Swann, G. M. P., M. Prevezer and D. Stout (1998), *The Dynamics of Industrial Clustering – International Comparisons in Computing and Biotechnology*. Oxford University Press: New York.

Tappi, D. (2005), 'Clusters, adaptation and extroversion – a cognitive and entrepreneurial analysis of the Marche music cluster,' *European Urban and Regional Studies*, **12**, 289–307.

Tichy, G. (2001), 'Regionale Kompetenzzyklen - Zur Bedeutung von Produktlebenszyklus – und Clusteransätzen im regionalen Kontext,' *Zeitschrift für Wirtschaftsgeographie*, **45**, 181–201.

Van Klink, A. and P. De Langen (2001), 'Cycles in industrial clusters: the case of the shipbuilding industry in the Northern Netherlands,' *Tijdschrift Voor Economische En Sociale Geografie*, **92**, 449–463.

Vernon, R. (1966), 'International investment and international trade in product cycle,' *Quarterly Journal of Economics*, **80**, 190–207.

Wuyts, S., M. G. Colombo, S. Dutta and B. Nooteboom (2005), 'Empirical tests of optimal cognitive distance,' *Journal of Economic Behavior & Organization*, **58**, 277–302.

Zeller, C. (2004), 'North Atlantic innovative relations of Swiss pharmaceuticals and the proximities with regional biotech arenas,' *Economic Geography*, **80**, 83–111.

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