Perspectives on Cluster Evolution: Critical Review and Future Research Issues

Michaela Trippl (michaela.trippl@circle.lu.se)
CIRCLE, Lund University

Markus Grillitsch (markus.grillitsch@circle.lu.se)
CIRCLE, Lund University

Arne Isaksen (anne.isaksen@uia.no)
Department of Working Life and Innovation, University of Agder

Tanja Sinozic (tanja.sinozic@wu.ac.at)
Institute for the Environment and Regional Development, Vienna University of Economics and Business

This is a pre-print version of a paper that has been submitted for publication to a journal.

This version: July 2014
Perspectives on Cluster Evolution: Critical Review and Future Research Issues
Michaela Trippl, Markus Grillitsch, Arne Isaksen and Tanja Sinozic

Abstract

The past two decades have witnessed an ever-growing scholarly interest in regional clusters. The focus of research has mainly been on why clusters exist and what characteristics “functioning” clusters hold. Although the interest in more dynamic views on clusters is not new, in recent years, however, more attention has been paid to providing better explanations of how clusters change and develop over time, giving rise to an increasing popularity of different variants of the cluster life cycle approach. This article offers a critical review of various cluster life cycle models. We discuss the key ideas and arguments put forward by their main protagonists and we identify several shortcomings – such as the problematic predefinition of development phases, indifference to context-specific factors and neglect of multi-scalar impacts – that surround these models. Based on this critical assessment the article identifies several core issues for future research. In particular, we argue that there is a need to gain a better understanding of the context sensitivity of cluster evolution, to explore how cluster development paths are influenced by a multiplicity of factors and processes at various spatial scales and their interactions, and to investigate the role of human agents and to unravel how they shape the long-term development of regional clusters.

JEL codes: R10; R11; R50; R58; P48

Keywords: cluster evolution; life cycle approaches; context sensitivity; multi-scalar frameworks; human agency

Disclaimer: All the opinions expressed in this paper are the responsibility of the individual author or authors and do not necessarily represent the views of other CIRCLE researchers.
Perspectives on Cluster Evolution: Critical Review and Future Research Issues


[1] CIRCLE, Lund University, Sweden, email: michaela.tripl@circle.lu.se
[2] CIRCLE, Lund University, Sweden, email: markus.grillitsch@circle.lu.se
[3] Department of Working Life and Innovation, University of Agder, Norway, email: arne.isaksen@uia.no
[4] Institute for the Environment and Regional Development, Vienna University of Economics and Business, Austria, tanja.sinozic@wu.ac.at

Abstract

The past two decades have witnessed an ever-growing scholarly interest in regional clusters. The focus of research has mainly been on why clusters exist and what characteristics “functioning” clusters hold. Although the interest in more dynamic views on clusters is not new, in recent years, however, more attention has been paid to providing better explanations of how clusters change and develop over time, giving rise to an increasing popularity of different variants of the cluster life cycle approach. This article offers a critical review of various cluster life cycle models. We discuss the key ideas and arguments put forward by their main protagonists and we identify several shortcomings – such as the problematic predefinition of development phases, indifference to context-specific factors and neglect of multi-scalar impacts – that surround these models. Based on this critical assessment the article identifies several core issues for future research. In particular, we argue that there is a need to gain a better understanding of the context sensitivity of cluster evolution, to explore how cluster development paths are influenced by a multiplicity of factors and processes at various spatial scales and their interactions, and to investigate the role of human agents and to unravel how they shape the long-term development of regional clusters.

Key words: cluster evolution; life cycle approaches; context sensitivity; multi-scalar frameworks; human agency

Word count: 9489
1 Introduction

Over the past two decades or so, there has been an enormous scholarly and policy interest in regional clusters. While the term ‘cluster’ was introduced by Michael Porter in the 1990s (Porter, 1990, 1998), the origins of the notion can be traced back to Marshall’s (1920) influential work on industrial districts. Nowadays, clusters have become an essential ingredient of regional development and innovation strategies in many parts of the world, and notwithstanding a number of critical evaluations (see, for instance, Martin & Sunley, 2003; Frenken et al., 2014), the notion of clusters is widely used in economic geography, innovation studies and related disciplines.

An extensive body of work has focused attention on explaining why clusters exist and what the main characteristics of “functioning” or fully developed clusters are. Compared to these static approaches, dynamic views on long-term cluster evolution have received less attention. Arguably, the question of how clusters develop and change over time is not new and has been a subject of research in the past (see, for instance, Storper & Walker 1989). In recent years there has been a growing recognition of the need to further develop and elaborate on dynamic perspectives on clusters to gain better insights into potential forms of their long-term evolution (see, for instance, Lorenzen, 2005; Bergman, 2008; Menzel & Fornahl, 2010; Isaksen, 2011). Popular approaches in this field of research are variants of the cluster life cycle (CLC) approach. Some of them have been heavily inspired by early work on product life cycles (Veron, 1966; Cox, 1967) and studies of industry life cycles (Klepper, 1997).

CLC models have essentially enhanced our understanding of crucial factors that may trigger the rise and further development of clusters. Life cycle approaches to cluster development, however, suffer from several shortcomings, most notably from rather deterministic views (Martin & Sunley, 2011) that preclude the capturing of the complexity and variety of cluster transformation evident from empirically grounded contextualised studies. The aim of this article is to critically review various CLC models and to identify a set of core issues that deserve due attention in future work on long-term cluster development. It is argued in this paper that, amongst other things, a stronger emphasis on the context sensitivity and multiscalearity of cluster evolution is necessary and that the role of human agents in cluster development needs to be further examined. We claim that consideration of these issues might enrich our explanations of how clusters evolve over time.

The remainder of the article is organized as follows. Section 2 introduces old and newer versions of the CLC concept that has gained increasing popularity in the literature on cluster development. We discuss the conceptual arguments and insights that have been provided by scholarly work on various CLC models. In Section 3 we identify a set of limitations that surround these approaches. Section 4 elaborates on an agenda for future research activities that may provide promising further insights into how clusters develop over time. Finally, Section 5 summarises our main arguments.

2 Cluster Life Cycle Approaches

Cluster life cycle (CLC) approaches depart from the idea that regional clusters go through different phases, often described as emergence, growth, sustainment, decline, and possibly renewal. In this section, we will first account for early work on the CLC approach by elaborating on Storper and Walker’s (1989) theory of geographic industrialisation. Then, we
will acknowledge the development of the CLC approach in the 1990s and 2000s before focussing on the more recent CLC models.

### 2.1 Early accounts of the cluster life cycle model

A discussion of the CLC approach could well start with the theory of geographical industrialization put forward by Storper and Walker (1989), as this is an early and influential account of important aspects of cluster life cycles. Geographical industrialization has in the words of Storper and Walker (1989, p. 9) four principal moments: localization, clustering, dispersal, and shifting centres. With their theory the authors aimed to bring growth and change to the forefront in studies of industrialization compared to the more static location theories.

The localization moment underlines that new or renewed, fast-growing industries have a large degree of locational freedom, denoted as windows of locational opportunity. Such industries can settle in a wide variety of locations. The specific places where innovating firms spring up are often chosen by the initial innovators. The locational freedom resembles Krugman’s (1991, p. 35) observation that highly localized industries “can be traced back to some seemingly trivial historical accidents”. Two main factors explain the locational freedom. First, new and renewed industries often demand new skills and work habits, and new types of input factors, that is, such industries are not bounded to location factors found in the old industrial centres. Second, fast-rising industries have high profits, “which allow them to attract the resources and labour that cannot be created at the new site” (Storper & Walker, 1989, p. 74).

The clustering phase focuses on how some of the places where new industries initially arise develop into expanding clusters. These centres benefit from increasing returns, enhanced technological capacity, capital and labour in-migration, and cumulative investments. Fast-growing industries thus produce their own locational specifications, and when some centres get well ahead of others, the window of locational opportunity shuts. The dispersal of industries from the dominant clusters takes place at some point of time. Often firms’ home base activities (Porter, 1998), such as top management, R&D, design, advanced product and service activities, remain in the clusters while more standardized activities spring up in “cost-cutting peripheries” (Storper & Walker, 1989). But the theory argues that clustered industries first of all spin-off growth peripheries; firms build new advanced plants in new territories to capture new markets. Eventually, dominant clusters in an industry may melt away. This is explained by the restructuring and renewal of an industry which opens up the locational window. Radically restructured industries usually shift to new locations, often outside previously dominant clusters, which may, for example, have an inflexible, unionized working class.

The theory of geographical industrialization predates much theorization of cluster life cycles. However, it appears to be quite US centred in reflecting large mobility of capital and workers in the US, especially the shift of the industrial core from the traditional manufacturing belt in the North-East to sunbelt areas (Norton & Rees, 1979).
2.2 Development of cluster life cycle approaches in the 1990s and 2000s

Since the work of Storper and Walker in 1989, the CLC approach has received considerable attention (for an overview, see Bergman, 2008). A central hypothesis of the life cycle approach is that clusters change with discernable phases of growth (Swann 1998; Braunerhjelm & Feldman, 2006). Each phase is characterised by specific features influencing shifts that are assumed to be generalizable across cluster populations. A broad starting point for explaining cluster formation and change is that clusters take off as Marshallian localised economies of scale and spillovers gain momentum, increase firm profits and create broader favourable business conditions that facilitate spin-offs and motivate firms to cluster together (Maskell & Kebir, 2006). Over time, the literature has extended this understanding of dynamics and found that clusters exhibit different likelihoods of stages depending on locational characteristics, technology area, sectoral specificities, the time period under investigation, and the market (local or global) of the cluster (Bergman, 2008).

A broad distinction can be made in approaches that focus on industry-driven explanations and such that emphasise processes specific to clusters (Martin & Sunley, 2011). Industry-driven explanations relate the growth of clusters to development stages in specific industries and technologies. In general, the importance of clusters is thought to be highest in early development phases of an industry and technology, when a lot of experimentation takes place and knowledge is not yet codified and standardised (Audretsch & Feldman, 2006; Ter Wal & Boschma, 2011; see section 2.3). “Cluster-specific” cycle views (Pouder & St. John, 1996; Iammarino & McCann, 2006; Maskell & Malmberg, 2007) suggest that cluster can grow or decline independently of the development of the industry, for reasons such as homogeneity or heterogeneity in competencies, and cluster-specific technological or institutional lock-ins.

In this section we try to synthesise the most salient features of cluster cycles that have been discussed in pioneering work. In doing so, we compare and contrast factors that characterise cluster stages. Many CLC studies have addressed the beginnings of cluster formation, that is, the emergence phase (see, for example, Braunerhjelm & Feldman, 2006). Cluster formation is partially caused by what has happened in the region (and globally) in the past, and the knowledge, problem-solving capacities and structures that are already available (Maskell & Malmberg, 2007). Triggering events for cluster emergence may have a partially random or serendipitous character, making it difficult to predict when and where clusters will arise (Braunerhjelm & Feldman, 2006). Cluster beginnings also differ according to the knowledge industries draw upon and contribute to. Science-based clusters have, for example, been traced back to discoveries generated in research universities, the continuous flow of skilled graduates sustaining a rich labour pool, as well as an entrepreneurial environment that surrounds the university and creates conditions for a cluster to emerge (Patton & Kenney, 2010).

As the cluster expands in size it enters the second stage (growth). This phase is characterised by a strong growth of leading firms and the entry of new firms. As knowledge exploration and application processes are not yet routinized, clustering is favourable for firm innovation processes because they lower search costs (both spatially and cognitively) (Maskell & Malmberg, 2007). The momentum of the growth stage is dependent upon alignment between local conditions such as skilled labour, training, suppliers, and an institutional framework that supports the exploratory processes and periodic adjustments to products and services that the firms are engaging in during this period (Bergman, 2008). Pouder and St. John (1996) argue that firms in clusters (or in their words “hot spots”) behave differently and are more innovative because of collective learning, enhanced legitimacy and agglomeration economies.
This phase is often connected to the life cycle of the industry (Audretsch & Feldman, 1996), where clustering may be favourable for innovations, but later as processes become routinized and other ways of organizing production become more efficient, the industry may become increasingly dispersed.

In the third phase, the cluster enters a period of maturation or exhaustion, where its processes and conditions no longer create products that are innovative or competitive, congestion costs rise, processes become routinized and easier to replicate elsewhere, and other ways of organizing production become more efficient, creating conditions for firms to leave the cluster (Swann, 1998). The number of firms falls, and existing networks become less fruitful sources of external information (Tichy, 1998, 2001). Cognitive bias, mimetic behaviour and institutional isomorphism lead to lock-ins and reduced innovativeness of cluster firms (Pouder & St. John, 1996). Clusters decline when entrepreneurial forces are stymied by a technological regime. A routinized regime is a barrier to entrepreneurial initiative because firm routines give them an advantage over the entrepreneur’s idea, restricting change.

Some clusters are able to enter a fourth phase, and reuse some skills and existing infrastructures and incrementally build novel industrial and sectoral identities. Bergman (2008) identifies amongst other things heterogeneity in agents and technologies as factors that may cause cluster rebirth. For example, if different actors are present the likelihood of exploiting different technologies to initiate a new cluster is higher.

2.3 Recent versions of the cluster life cycle approach

While clusters have been debated for more than two decades, literature on evolutionary economic geography has prospered more recently (Boschma & Frenken, 2006; Essletzbichler & Rigby, 2007; Martin & Sunley, 2007; Boschma & Martin, 2010). The latter literature emphasizes that economic evolution in a geographic context depends on factors such as heterogeneity of firm competencies, related variety, the evolution of networks, and path-dependency. Recent versions of the CLC approach have partly taken up these factors.

Firm heterogeneity and variety feature prominently in the models developed by Menzel and Fornahl (2010) and Ter Wal and Boschma (2011). Menzel and Fornahl (2010) propose a clear distinction between cluster firms, firms in the same industry located elsewhere, and firms in other industries but located in the same region. While appreciating the role of interactions between these different types of firms, the institutional context and the industry life cycle, it is argued that firm heterogeneity and localised learning processes are the central factors explaining cluster change. Menzel and Fornahl (2010) suggest that the development of technological relatedness between firms is a precondition for the emergence of a cluster while heterogeneity is considered as crucial source for the extension or renewal of development trajectories. Clusters begin “…in those regions where the knowledge bases of companies converge around technological focal points” (p. 231). Technological convergence underlying the momentum of cluster formation is shaped by, amongst other factors, interactive learning processes between heterogeneous firms in geographic proximity to one another. Firm heterogeneity can be increased through learning with non-cluster firms both locally and globally. This may bring in new knowledge to the cluster, shifting its thematic boundaries. Menzel and Fornahl’s model suggests that localised learning dynamics and firm heterogeneity propel clusters through life cycles. Despite this dominant trajectory, the authors also open up for alternative trajectories. For instance, without technological convergence in the emergence stage, a cluster may never reach the growth stage. Also, by introducing heterogeneity in later
stages, clusters can continuously renew themselves and do not necessarily need to decline. In sum, “clusters display long-term growth if they are able to maintain their diversity” (Menzel & Fornahl, 2010, p. 218).

Ter Wal and Boschma (2011) propose a framework in which clusters co-evolve with firm capabilities, industry life cycles and networks. The authors emphasise the importance of variety as regards firm capabilities that resonates well with the model introduced by Menzel and Fornahl (2010). In addition, the framework of Ter Wal and Boschma (2011) elaborates on the effects of networks for the evolution of clusters. Ter Wal and Boschma (2011) argue that networks and firm capabilities are interrelated so that “variety across firms in terms of capabilities drives the evolution of networks through time” (p. 923). Firms with strong capabilities will be attractive collaboration partners and thus more centrally positioned in networks. However, due to the high degree of uncertainty during cluster emergence as regards the future dominant technologies and players, networks are highly unstable. Firms are expected to switch network partners frequently, whereas the choice of partners depends both on social networks and chance events.

As clusters grow, several forces lead to stable core-periphery network patterns. This implies first that firms with superior capabilities are the most attractive network partners and thus are centrally positioned. The central network position further increases the attractiveness of these firms. Preferential attachment is stimulated by the advantageous network position of pioneers, the higher likelihood of firms in weaker positions to exit the industry, the importance of prior alliances for the formation of new ones, as well as the technological trajectory where the centrally positioned firms with superior capabilities are likely to further drive technological development. As the technology has not yet matured, tacit knowledge plays an important role, which in consequence stimulates networks in geographic proximity for reasons such as social capital, trust and the ease of face-to-face interactions. During maturity and decline, many firms exit the industry. Firms at the core of the network (and frequently located in industry clusters) tend to have a higher likelihood to survive. The endurance of networks can have, however, distinct disadvantages because of a decreasing variety of firm capabilities, which may lead to cognitive lock-in, and an increasing codification of knowledge, which reduces the need for geographic proximity. Nevertheless, a new cluster life cycle may be started if cluster firms succeed in generating a new technological breakthrough. However, similar as in the introductory stage, such technological breakthroughs will often be generated outside the cluster leading to significant changes in the network structure (Ter Wal & Boschma, 2011).

Martin and Sunley (2011) emphasize the unpredictable nature of development paths, which implies that clusters may follow a variety of potential trajectories. This is in line with the path dependence theory in evolutionary economic geography, where path dependence does not imply historical determinism but is seen in relation to mechanisms propelling path creation and path destruction (Martin & Sunley, 2006; Simmie, 2012). Martin & Sunley (2011) emphasize that the trajectories of clusters are unpredictable, mainly because they consist of agents who learn, interact and respond to their perceptions about the current state and future development within the cluster and their environment. The authors propose to think about clusters as complex adaptive systems and apply an adaptive cycle model. In contrast to the traditional approaches, the adaptive system model allows for a variety of development trajectories. In their slightly modified version, Martin & Sunley (2011) identify six possible evolutionary trajectories. One follows the typical life cycle of emergence, growth, maturation, decline and eventual replacement. The notion of replacement strongly builds on the idea that existing resources are released and brought to new use. The other possible trajectories include
different combinations of cycles. Thus, clusters do not necessarily need to move from a growth to maturity stage. Particularly clusters with strengths in generic or general-purpose technologies, usually associated with high-tech industries, may continuously innovate and mutate. The heterogeneity of firms remains high due to ongoing intensive innovation activities. This is often linked to geographically open knowledge networks. While there might be a high connectedness within the cluster, firms have established interregional, sometimes global, linkages, a feature that has been observed for high-tech clusters. Also, the future technological paths remain uncertain. This in turn requires a strong endowment with venture capital so that firms can embark in uncertain, radical innovation activities. Such clusters keep a high degree of resilience. In addition, Martin and Sunley (2011) illustrate that clusters may for instance fail to grow, be replaced and disappear, or stabilise after the growth phase even in mature industries.

To summarize, the review of the literature on cluster life cycles has shown that since Storper & Walker’s (1989) theory of geographical industrialization a rich body of conceptual work has emerged, seeking to shed light on how clusters develop over time. Life cycle approaches to cluster development, however, have recently come under increasing critical examination. In the next section we identify some key limitations of these approaches.

3 Shortcomings of the CLC Approaches

The CLC models discussed in the previous section are not uncontested. In this section we present their main criticism, focusing in particular on the rather deterministic view of cluster development, derivation of cluster life cycles from industry life cycles, lack of (regional) context sensitivity, neglect of multi-scalar impacts, and limited appreciation of the role of human actors. Before doing so, it is important to note that some of the shortcomings identified below hold only true for some CLC models but less so for others.

Various authors (see, for instance, Martin & Sunley, 2011; Oinas et al., 2013) have criticized the deterministic logic of life cycle approaches to cluster development, which carries biological connotations and imply “… some sort of ‘aging’ process. But in what sense can clusters be thought of as having ‘lives’ or ‘ageing’ or passing through ‘life stages’?" (Martin & Sunley 2011, p. 1300). Adopting a life cycle approach implies that cluster development follows a predetermined sequence of stages from birth to growth, maturation and decline (see, for instance, the model proposed by Poudre & St. John, 1996) and possible renewal (that is, the start of a new cycle). Also more recent work (Maskell & Kebir, 2006; Ter Wal & Boschma, 2011) takes the prescribed sequence of phases as granted. Such a view does not allow for capturing other development patterns such as a hyper growth stage after a growth stage, nor paths that never reach the subsequent stage (Bergman, 2008) and other forms of development patterns observable in the real world (Oinas et al., 2013).

Any CLC model that seeks to distinguish specific development stages contradicts the important idea of non-linearity in evolutionary thinking (see, for instance, Dosi, 1991). To take the evolutionary approach seriously means that we cannot fully foresee the future evolution of any cluster. As Sydow et al. (2012, p. 159) remind us, “it makes sense to speak of a path only in those instances where competing options existed and the later ‘solution’ was not foreseeable at the beginning of a path”. To follow evolutionary thinking then means that
the development of a cluster can be explained only ex post but not ex ante since the future course of a path is open.\footnote{We are thankful to an anonymous reviewer of a previous version of this paper for valuable comments that helped us to elaborate on and sharpen our argument on this issue.}

As noted above, CLC models postulate that clusters follow some distinct phases; emergence is followed by growth, then maturity, and so on. However, such a logic (that is, taking the existence of a predetermined development pattern for granted) is not applied to explain the rise of clusters, or to answer the question of where clusters in new industries first make their appearance. Some CLC approaches consider the actual location of new clusters as the result of historical accidents, suggesting that new clusters often “start out in a particular location more or less by chance” (Maskell & Malmberg 2007, p. 612). It may seem inconsistent that clusters first emerge randomly in specific places but then follow a predefined development logic when they have reached a certain size or start growing. Also, according to evolutionary thinking cluster development can be understood quite differently. The emergence of clusters can be based on previously developed local capabilities, routines and institutions (Boschma & Frenken, 2011). “Even in the case of radical technological development, knowledge production is also highly cumulative and builds on pre-existing localized scientific and technological resources” (Tanner, 2011, p. 24). After their emergence clusters can take very different paths that may also differ from the canonical routes suggested by various CLC approaches. Thus, “path dependence and path creation are only two possible ways to build and transform a path in time and space; others are intentional path defence or extension, unintended path dissolution, or breaking a path without creating a new one” (Sydow et al., 2012, p. 158). This view allows for the fact that “history matters” but also that history does not entirely restrict possible development paths.

Some CLC models claim that the development (or more precisely, the life cycle) of a cluster is tightly linked to or can directly be derived from the life cycles of products, industries and technologies (Veron, 1966; Utterback & Abernathy, 1975; Klepper, 1997) that ‘drive’ the cluster. Other protagonists of the CLC approach, in contrast, convincingly argue that the development patterns of cluster and industries may be different (Menzel & Fornahl, 2010). Clusters in specific industries are certainly affected by general market and technological developments within these industries. Industrial classifications may, however, include rather diverse activities. Different parts of industries may display different development dynamics, and clusters in similar industries may also reveal quite different forms of path development. Belussi and Sedita (2009) demonstrate that firms in one footwear district in Italy diversified their products, which attracted some big luxury brands to the district, while firms in another Italian footwear district carried out a cost-led strategy and outsourced production to low-cost countries. This demonstrates that no one-to-one relationship exists between industry characteristics and cluster life cycles and that, amongst others, context-specific regional factors are neglected in CLC models.

Further, CLC models tend to give primary emphasis to a limited set of influences and tend to ignore that cluster evolution is based on multiple factors (Martin & Sunley, 2011). Some of the recently introduced CLC models, for example, focus on characteristics and dynamics of firms, their capabilities and networks, whilst regional characteristics (and their interrelationships with firm characteristics and extra-regional factors) are not on their agenda. The literature on cluster life cycles has thus far devoted relatively little attention to how cluster development paths are shaped by a multiplicity of factors at various spatial scales and their interaction. It is fair to argue that the multiscalarity of cluster evolution remains poorly.
understood. For example, the relation between various configurations of knowledge linkages at different geographical levels and cluster evolution has not been a key topic of interest in the CLC literature and little has been said about the insertion of clusters in institutional frameworks, which are multiscalar in nature.

CLC models can also be charged as being too much occupied with the structure level, such as firm structure (heterogeneity) and knowledge networks, while agents and their activities are not sufficiently taken into account. This critique relates to Martin and Sunley (2006) who have posed the question whether path dependence does need a theory of human agency. “In what ways is path dependence intentionally created by actors, or an unintentional emergent effect at system level?” (op. cit., p. 404). Maryann Feldman, in particular, has brought human agency in the form of entrepreneurship into research of cluster evolution. The importance of entrepreneurs is analysed both from a conceptual and empirical point of view. In conceptual terms entrepreneurs are seen to “spark cluster formation and regional competitive advantage” (Feldman et al., 2005, p. 130) and also to “build resources and community” (op. cit., p. 131). The rather low attention CLC models have given thus far on agency may be explained by the fact that CLC studies mainly include empirical investigations of specific phases in the development of clusters and have rarely examined individual clusters and the importance of shifting actors for a long time. Two widely recognized edited volumes, for example, concentrate on cluster genesis (Braunerhjelm & Feldman, 2006) and on emerging clusters (Fornahl et al., 2010).

Finally, two further limitations of CLC approaches are worth to mentioning briefly. First, CLC models have been criticized for not offering enough theoretical explanations of the causes for which clusters switch from one stage to another (Boschma & Fornahl, 2011; Oinas et al., 2013). Second, the empirical foundation of CLC approaches seems to be rather weak. So far, the life cycle of clusters has mainly be based on a review of unrelated empirical studies of clusters of different age, in technological areas, and so on but empirical tests, using the case of a single cluster, are rare or even missing.

### 4 Key Issues for Future Research

This section identifies some key challenges for future research activities on long-term cluster development. There are still many unresolved questions of how clusters evolve over time (see, for instance, Boschma & Fornahl, 2011). We do not intend to construct an overarching research strategy or map out a comprehensive research agenda here. This would be far beyond the scope of this paper. Instead we focus on three core issues that in our view should receive due attention in future research on cluster evolution.

#### 4.1 Context sensitivity

As noted in section 3, there are strong reasons to assume that there is more than one potential development path of cluster evolution. The multiplicity of possible development paths can partly be explained by the context sensitivity of clusters. A dominant strand of the CLC literature links the development of clusters to the development of industries and technologies while other CLC approaches focus on cluster internal factors that are presumed to explain the

---

2 We are very grateful to an anonymous reviewer of a previous version of this manuscript for raising this point.
trajectories of clusters (Martin & Sunley, 2011, see also section 2). Hence, there is some tradition in the CLC literature to consider the sectoral and technological context. However, also the regional and national context is of importance for the development of industries and clusters as argued in the literature on national innovation systems (Lundvall, 1992; Nelson, 1993, Edquist, 2005), varieties of capitalism (Hall and Soskice, 2001) and regional innovation systems (Cooke, 2001; Asheim and Isaksen, 2002, Tödtling & Trippl, 2005). Furthermore, also the nature of knowledge bases (Asheim & Gertler, 2005) and their combination (Asheim et al., 2011) may have an influence on how clusters develop over time. In the following we focus on the latter two notions, that is regional innovation systems and knowledge bases, and we offer some brief thoughts on how these concepts could be employed to increase our understanding of cluster evolution as a context-specific phenomenon.

The Regional Innovation System (RIS) concept devotes attention to the companies, cluster structures, knowledge providers and the institutional set-up of a region, as well as to knowledge connections within the region and to the external world. The region is seen as a crucial level at which innovation is generated through knowledge linkages, clusters and the cross-fertilising effects of research organisations (Asheim & Gertler, 2005). RISs come in many shapes and there are strong reasons to assume that cluster development is linked to the configuration of RISs. Cooke (2004) distinguishes between entrepreneurial and institutional RISs and claims that the former offer excellent conditions for the development of high-tech clusters, whilst the latter provide a fertile ground for the evolution of traditional ones. Other scholars contrast organizationally thick RISs with thin ones and provide empirical evidence that clusters operating in the same industry but located in these different types of RISs display diverging development and innovation dynamics (Chaminade, 2011; Tödtling et al., 2011, 2012). These typologies are useful for analysing how the evolution of a cluster might be influenced by the RIS in which it is embedded.

The relation between RISs and cluster evolution is complex. According to the RIS approach, clusters form an integral part of RISs. Empirical studies on RISs provide evidence that the emergence, growth, maturity, decline and possibly renewal of clusters can, thus, only be understood if the specificities of the knowledge infrastructure, institutional set up, cultural aspects and policy actions of a particular region are considered. For instance, RISs that already host dynamic high tech clusters provide favourable conditions for the emergence of new ones, even if these newly emerging clusters are different from those developed earlier. Such RISs offer essential conditions, such as excellent research institutes, venture capitalists, a pool of highly skilled mobile workers and dense communication networks (see, for instance, Prevezer, 2001). RIS that are poorly endowed with such structures, experiences, and knowledge assets are likely to take different routes. The rise of new (high-tech) clusters in such regions is less a spontaneous phenomenon but depend more on the inflow of external knowledge, expertise and market intelligence and a stronger role of policy (Leibovitz, 2004). In addition, new cluster formation in such regions is inextricably linked to a transformation of the RIS that becomes manifest in the creation of a variety of new organisations, processes of institutional (un)learning and socio-cultural shifts. Another example is the literature on the renewal of traditional clusters. Much of this work is focused on old industrial regions, emphasizing a strong relationship between the rejuvenation of mature clusters and prevailing RIS characteristics including their transformative capacities. Conceptual and empirical research on this issue suggests that the presence or absence of favourable RIS structures and RIS changes can make a difference, influencing the success of regeneration processes of traditional clusters (Trippl & Otto, 2009; Hassink, 2010).
The trajectories of clusters are also dependent on the knowledge bases available in the RIS. The knowledge base concept has enhanced the understanding of sectorial variations that may exist within a RIS and explains how they are related to different types of innovations, knowledge dynamics and learning activities. The knowledge base concept argues that firms, industries and clusters can be classified based on the type of knowledge that is critical for innovation (Asheim & Gertler, 2005; Asheim et al., 2011). It distinguishes in synthetic, analytical and symbolic knowledge. An analytical knowledge base prevails in high-tech clusters such as biotechnology or nanotechnology where innovation is driven by scientific progress. A synthetic knowledge base is dominant in mature clusters operating in fields such as industrial machinery or food processing. Innovation is based on the use and new combination of existing knowledge and learning by doing, using and interacting. The symbolic knowledge base is present in creative and cultural clusters (advertisement, fashion, new media and design). Innovation is devoted to the creation of aesthetic value and images. Recent conceptual work suggests that the combination of different types of knowledge bases within a RIS (Asheim et al., 2011) can play an important role for the emergence of new clusters or the renewal of existing ones. Conducting thorough empirical investigations of the nexus between combinations of knowledge bases and cluster development may be a promising avenue for future research.

Overall, there are convincing theories, concepts and some empirical evidence that the development paths of clusters depend upon RIS architectures and knowledge bases. Typologies proposed in the RIS literature can help to understand why and under which conditions clusters take certain trajectories, sharpening our view of the context-sensitive nature of cluster evolution. However, as argued at the beginning of this sub-section, contextual conditions at other spatial scales and related to sectors and technologies also matter. Analysing how important contextual factors influence cluster trajectories can enhance the understanding of cluster evolution as a context-specific phenomenon.

4.2 A multi-scalar perspective on cluster evolution

A challenge for future research on cluster evolution is to further our understanding of how factors at different geographic scales and analytical dimensions interact and influence cluster development paths. Martin & Sunley (2011) suggest that the evolution of clusters depends upon multiple factors and that CLC approaches pose a risk of neglecting some important ones. On the other hand, different CLC concepts emphasise different factors. For instance, the more recent CLC approaches introduce a focus on firm heterogeneity and networks. Often CLC approaches interpret the trajectory of clusters in relation to the development of an industry or a technology. As argued in section 4.1, the trajectories of clusters further depend upon their embeddedness in a spatial context. How all these factors are interrelated remains an open question that deserves due attention in future research.

The importance of different spatial scales for the evolution of clusters becomes apparent in various strands of the literature. One important aspect relates to the scale of knowledge linkages of firms. There is strong empirical evidence and conceptual underpinning that the combination of regional and interregional knowledge linkages is most conducive to the innovativeness and competitiveness of firms and clusters (see, for instance, Bathelt et al., 2004; Wolfe & Gertler, 2004). Clusters are likely to differ in their evolution according to their degree of embeddedness in multi-scalar networks and the degree of complexity resulting partly thereof (Bergman, 2008; Martin & Sunley, 2007). A cluster that is similar to a complex system is likely to evolve in a non-linear manner with multiple feedback mechanisms between...
its components (Martin & Sunley, 2007), the influences of which are likely to change over time.

Furthermore, there is evidence that the embeddedness of clusters in local, national and supranational institutional contexts has a strong impact on their development (Gertler, 2010; Grillitsch, 2014). The literature on varieties of capitalism suggests that the national institutional configurations lead to competitive advantages in particular industries (Hall and Soskice, 2001). On the other hand, Strambach (2010) argues that specific clusters and industries can emerge rather surprisingly despite an unfavourable national institutional environment. Local institutions related to the cluster or in broader terms related to the RIS in which clusters are embedded also matter. For instance, various studies have found local traditions, values and social proximity to be a defining feature of a cluster’s evolution (Saxenian, 1994; Sydow & Staber, 2002). However, as Gertler (2010, p. 6) states it, “one needs to understand far better than is done currently exactly how institutional forms and the incentives they create at any one particular scale influence, are influenced by, and interact with, the institutional architectures that are erected at other geographical scales”.

A possible approach for assessing interdependencies across multiple scales and analytical dimensions is to adopt a micro-meso-macro framework. Dopfer et al. (2004) propose for instance a framework where “micro” relates to the characteristics and behaviour of individual agents and firms while the meso-level relates to the institutions, in which agents’ behaviour is embedded. The “macro” relates to the structure between the meso-levels. The interactions between the micro, meso and macro levels come into focus. Also, Martin and Sunley (2011, 2012) explore the distinction between micro and macro. Similar to Dopfer et al. (2004), “micro” is referred to as the behaviours and interactions of the components of a system, that is, firms or individuals. “Macro” applies to cluster-wide structures, which can relate for instance to cluster-specific institutions or organisations. Martin and Sunley (2012) point to a bi-directional causality where firms or individuals can influence the macro structures (upward causation) while at the same time macro structures such as institutions influence the behaviour of firms and individuals.

While the discussion on micro (meso) macro is theoretically intriguing, it remains to be seen how these interdependencies materialise in specific empirical settings. For instance, possible questions are how important upward or downward causation is in specific empirical settings, specific stages of cluster development, or how this differs by types of clusters. Also, there remain many open questions regarding conceptualisations of the meso level. If it is interpreted as cluster-specific institutions and organisations, how can the meso level be interpreted in the wider context a cluster is embedded in, such as the RIS, the national context, global innovation networks, or the sector and technology? Hence, a multi-scalar approach to cluster evolution will be required.

4.3 The role of human agency

As discussed in section 3, CLC models can be criticized for paying insufficient attention to how agents and their activities influence the evolution of clusters. If one adopts a critical realist approach (Sayer 1992) the development of a specific cluster can be analyzed by studying 1) its structure and composition, and 2) specific conditions and events. Some structural characteristics may be necessary for the evolution of, for example, a biotechnology cluster, such as access to specialized scientific knowledge. Such knowledge may be available in several places. However it is only “activated” in those areas where entrepreneurs manage to
commercialize the knowledge with the support of venture capitalists. Thus, the evolution of a cluster then depends upon “conditions whose presence and configurations are contingent” (Sayer, 1992, p. 107), that is, found at certain places and times. The argument here is that one cannot explain the evolution of a cluster by focusing only on its structural characteristics, industry composition and so on. The activities of agents, such as the decisions by entrepreneurs or firms (where) to establish, expand or relocate their business, are also critical for the development of a specific cluster.

Some scholars argue that the renewal of old clusters, for example, presupposes a well-developed RIS (Trippl and Otto, 2009; see also section 4.1). However, it also depends upon the formation of new innovation networks and new policy approaches such as the active support of university-industry links (op. cit.). Many old industrial clusters may have strong RISs, which may also be a reason for lock-in and stagnation. Strong RISs and new innovation networks and policy approaches are critical factors for the rejuvenation of old clusters, according to Trippl and Otto (2009). Although strong RISs are socially constructed, the core argument here is that new networks and new policy approaches demand human agency, for example by firm leaders who find new innovation partners and policy makers or politicians who initiate new policy instruments or strategies. Understanding the key role that can be played by human agents in cluster evolution is, indeed, a key challenge for future research.

We will underline that human agency must be understood broadly to include actors at different geographical levels and in both the public and private sector. This is clearly emphasized by Mazzucato (2013) who stresses the fundamental role of the state, through its various agencies and laboratories, in the development of radical technological innovations, and in the creation of new research based clusters. She shows that in particular government labs and government-backed universities, which are often involved in large, mission-oriented research programs in the US, have been crucial for producing radical new technologies and products. The state is presented as an entrepreneurial agent, more precisely the politicians and bureaucrats who allocate resources to large-scale, risky and uncertain research projects, and the large number of researchers and research teams that engage in early radical research. The state has also acted as a market shaper by creating networks between state agencies and private sector actors that facilitate commercial development.

Although the US state (and other nations) has been very proactive and entrepreneurial in the development and commercialization of new technologies, private sector entrepreneurial activity is also vital. Government funded research programs require complementary assets in firms in order to commercialize and industrialize technological innovations that are the fundament for new clusters. Mazzucato (2013, p. 11) documents that “there is not a single key technology behind the iPhone that has not been State-funded”. But Apple gathered experts together that were able to integrate the technologies, provide aesthetic design, perform great marketing, and so on. And Steve Jobs’ “steadfast focus on architectural innovations that disrupt the markets in which they compete are the reasons why he managed and deserved to capture a significant share of the rewards” (op. cit., p. 175).

Mazzucato’s argument on the role of state financed radical research is mostly relevant for the development of science-based clusters. The argument focuses on the development of new to the world technology through mission oriented, government funded research and its commercialization often by private actors with high absorptive capacity. Clusters operating in traditional industries rely upon other types of human actors. However, also in this type of clusters, studies could adopt the approach suggested by Mazzacuto (2013) in which an
underlying and implicit message is “to follow the knowledge”. Possible issues on a research agenda could include where critical knowledge for the development of clusters comes from, who (which persons and organization) has created the knowledge, who has put it into use, who has profited from its use and so on. An important methodological approach in such studies is intensive research to map important actors and events for the development of clusters.

5 Summary

Over the past few years there has been a growing recognition of the need to move beyond static views in the cluster literature and focus more attention on how regional clusters develop over time. This shift towards a stronger interest in dynamic perspectives have been accompanied by and resulted in a growing popularity of various cluster life cycle models. However, as shown in this paper, the question of how clusters evolve is not new and application of life cycle ideas has a long history in cluster research. The paper has provided a critical review of old and newer life cycle approaches to cluster development. It has been demonstrated that since Storper and Walker’s (1989) well-known theorization of geographical industrialization, which can be regarded as an early and influential account of important aspects of cluster life cycles, various intellectual efforts have been made to conceptualize cluster evolution by applying life cycle ideas. These approaches, however, are not uncontested. We have identified several limitations, criticizing – amongst other things – the rather deterministic view that results from assuming a prescribed sequence of development phases and contradicts evolutionary thinking, under-appreciation of regional and other context-specific factors, ignorance of the role of human agents, and a limited consideration of multi-scalar influences. Partly based on this critical assessment we elaborated on several key issues that deserve due attention in future research on cluster evolution. We focused on three main tasks. First, we claimed that one of the key challenges is to gain more insights into the context specificity of cluster evolution by examining how different types of RISs and knowledge bases shape cluster trajectories. Second, we advanced the argument that a multi-scalar framework should be adopted to enhance our knowledge about how multiple factors at various geographical scales and their interdependencies influence cluster development. Third, we have identified the need to get a better understanding of the role of human agents and how long-term cluster evolution is shaped by their activities.

References


