Conceptualising Cluster Evolution: Beyond the Life-Cycle Model?

Ron Martin and Peter Sunley
Conceptualising Cluster Evolution: Beyond the Life-Cycle Model?

Ron Martin* and Peter Sunley**

*Professor of Economic Geography, University of Cambridge, UK (rlm1@cam.ac.uk)

**Professor of Human Geography, University of Southampton, UK (P.J.Sunley@soton.ac.uk)

Revised Draft: June 2011
Abstract

Although the literature on the evolution of industrial clusters is not vast, a preferred approach has already become evident, based around the idea of a cluster 'life-cycle'. This approach has several limitations. In this paper we explore a different conception of cluster evolution drawing on the 'adaptive cycle' model that has been developed in evolutionary ecology. Using this model, cluster evolution is viewed as an adaptive process with different possible outcomes based on episodic interactions of nested systems. Though not without limitations, this approach offers greater scope as a framework for shaping the research agenda into the evolution of clusters.

Key Words

Clusters   Evolution   Life-cycle model   Complex systems   Adaptive cycle model

JEL   R00 R1 R3
Introduction

It is now two decades since Michael Porter first introduced an explicit geographical dimension to his cluster idea (Porter, 1990), thereby instantly creating a spatial concept that has come to exercise extraordinary influence over both academic work and policy discourse. Clusters have become a key concept in economic geography, urban studies, regional economics and related disciplines; and policy-makers everywhere have seized on the notion as a tool for promoting regional growth and competitiveness. Notwithstanding some cautionary assessments of the concept - concerning its definition, empirical measurement, theorisation, claimed impacts on local innovation and productivity, and its problems as a policy tool (see, for example, Martin and Sunley, 2003; Asheim, Cooke and Martin, 2006; Martin and Sunley, 2011) - the cluster idea has achieved something akin to hegemonic status. And perhaps not surprisingly, since it seems to embrace - or rather is sufficiently flexible a concept that it can be made to embrace - a wide range of examples and types of localised economic specialisation, both past and present (Karlsson, 2008).

The reference to past examples is of more than incidental interest. For the economic landscape abounds with clusters and other types of industrial district and agglomeration that appear to be at different stages of development, or at least different historical stages and show different degrees of economic success. At one end of the spectrum are the remnants of clusters that were once thriving national or even international leaders in their particular industries, but which long ago lost that dynamism. Many are struggling to survive in a highly competitive global economy. (Some former clusters of course have not survived, and have long since disappeared altogether from the landscape). Then there are others that have shrunk in size and importance but seem to have managed to undergo reinvention or rejuvenation. Then at the other end of the spectrum are more recently developed, highly successful clusters that are playing a formative role in their respective regional and national economies, and indeed at the global scale. Yet further, there are others that are at an even earlier, embryonic stage of development, still in the process of formation. Clusters come and go; they
emerge, grow, may change in complexion and orientation, may undergo reinvention and transformation, and may eventually decline and even disappear. In short, they evolve.

Porter himself has not had that much to say about cluster evolution, beyond the brief discussion of some of the reasons for the birth, growth and decline of clusters in his book *On Competition* (1998, pp. 237-245). Certainly he has offered no coherent theory of cluster evolution. In recent years, however, there has been a growing interest by economic geographers and others in constructing just such a theory. To some extent this interest can be seen as being part of, or stimulated by, the emergence of an evolutionary paradigm in economic geography (Boschma and Frenken, 2006; Martin and Sunley, 2006; Boschma and Martin, 2007; Boschma and Martin, 2010) in which the focus is on understanding how the economic landscape, including clusters, evolves and transforms over time. The problem is how best to forge that understanding.

Although the literature is not vast, a preferred approach in cluster research has already become evident, based around the idea of a cluster ‘life-cycle’. Of course, the application of the ‘life-cycle’ idea to economic phenomena is hardly new: it can be traced back to the early-1950s, when it was introduced to explain the expected market and sales profile of a typical product over time, in terms of a four or five stage sequence. From there the idea was taken up to describe the development over time of firms, industries and technologies, even whole national economies (Storper, 1985). ‘Life-cycles’, it seems, are observable everywhere in the economy. It was inevitable, therefore, that the notion would be used to characterise the evolution of clusters, especially given that the latter are typically based on, or defined in terms of, a dominant industrial or technological specialism.

Our argument in this paper, however, is that the ‘life-cycle’ concept – though irresistibly heuristic as a metaphor – may have limits as a characterisation of how clusters evolve over time. The problem is not simply one of metaphor *per se*. The use of metaphors is all but inescapable in the conduct of socio-economic enquiry. What matters is the appropriateness and plausibility of the metaphor that is transferred from one disciplinary domain (from biology in
the case of the ‘life-cycle’) to another (in our case, industrial clusters). For metaphors are not just innocent short-hand descriptors: they act as models, in the sense that they typically embody, explicitly or implicitly, causal type arguments. The idea of the ‘life-cycle’, for example, implies some sort of ‘ageing’ process. But in what sense can we talk of clusters having ‘lives’ or ‘ageing’ or passing through ‘life-stages’? Metaphors can indeed be a source of inspiration and theoretical and explanatory advance (see Wimmer and Kössler, 2006). But to have this role, they must have a meaningful ontological ‘fit’ in the field into which they are being transferred.

This issue of ontology can be taken further. Can products, technologies, industries, and clusters be treated as if they are the economic equivalent of biological organisms? This may be a valid assumption; but it is certainly not the only possible conception of clusters. In fact, as we explore in this paper, an alternative way of thinking about clusters is that they are complex adaptive systems. If this perspective is adopted, it opens up a range of possible models for how clusters might evolve. We view this as a positive feature since we are sceptical of any suggestion or assumption that there exists one single model of cluster evolution. We know that for every instance of a product, industry or technology that is found to follow the ‘classic’ life-cycle pattern (‘birth’, ‘growth’, ‘maturity’, ‘decline’ and ‘death’), there is (at least) another that cannot be so readily depicted. Indeed, many product, technological or industrial developmental paths do not follow straightforward or recognisable life-cycle sequences. Since the ‘life-cycle’ model first appeared in industrial and business studies it has been repeatedly questioned on these grounds. Research in organisational science, for example, has sought to move beyond standard life-cycles models of business evolution to more nuanced perspectives capable of capturing the rich variety of firm and product evolution that is readily observable. And this questioning of the life-cycle framework surely carries over to clusters. Further, a complex adaptive systems approach raises questions about the importance of system definitions for the mental models we use to describe and order ideas about system change. This seems especially relevant in the case of clusters (see Martin and Sunley, 2003). In other words, how we think about cluster evolution also depends on how conceptualise clusters as systems. While ambiguities and limitations also
surround a complex adaptive systems approach, its value is that it forces us to think about how cluster identity relates to the problem of cluster evolution. Our paper thus falls into two main parts. We begin by highlighting some of the limitations of the basic ‘life-cycle’ model of cluster evolution, and how recent studies have sought to overcome them whilst still retaining the basic notion of the model. We then turn to explore the scope for and limits of a complex adaptive systems perspective.

The Life-Cycle Approach and Its Limits

According to Van de Ven and Poole (1995), life-cycle theory focuses on the development of a single organism and portrays change as imminent: that is, “the developing entity has within it an underlying form, logic, program, or code that regulates the process of change and moves the entity from a given point of departure towards a subsequent end that is prefigured in the present state” (p. 515; our emphasis). Thus external events influence the process but they are always mediated by the imminent internal logic, rule-set or programme that governs the entity’s development. The typical sequence of events is thus unitary, singular, cumulative and conjunctive as stages are derived from a common underlying process. In such a sequence “Each of these events contributes a piece to the final product; they must occur in a prescribed order, because each piece sets the stage for the next. Each stage of development is seen as a necessary precursor of succeeding stages” (Ibid, p. ??). A strict definition of the life-cycle thus implies both an imminent logic and necessary historical sequence.

The stylised life-cycle model of cluster development closely follows this assumed logic. Typically a cluster’s ‘life’ is portrayed as consisting of five main stages or phases - emergence or birth, growth, maturity, and decline (and even ‘death’)– usually delimited in terms of changes in some metric used to measure or proxy for the ‘age’ of the cluster, for example numbers employed, number of firms, innovativeness, market share, or some such similar indicator. Further, these stages are attributed to the unfolding of key systemic characteristics that act, in effect, as ‘ageing’ processes.
In the cluster literature, the rules governing cluster development and ‘ageing’ are either traced to an underlying industrial-technological cycle, or to processes and mechanisms specific to agglomerations. In the first approach, the cluster life-cycle is based on a one-to-one synchronicity between the cluster’s vitality and the \textit{stage reached by the technology and industry life-cycle}. That is, it is assumed that as the technology–industry life-cycle matures, cluster advantages gradually become disadvantages and industries tend to disperse spatially as process innovations become more important (Swann, 1998). There are several different versions of this industry-driven approach. Most argue that clusters are most economically valuable and beneficial in the early stages of an industry’s evolution (Audretsch and Feldman, 1996). Local external economies, it is argued, are especially important during the early stages of an industry and technology, so that clustered firms are said to outperform their non-clustered rivals at this time (Van Klink and De Langen, 2001).

However, clustering becomes less beneficial as a dominant product design emerges and the focus of innovation shifts to process and productive efficiency, rather than new products. Clustering then becomes disadvantageous to firm performance either because of high and rising congestion and transaction costs, or because of a collective ‘lock-in’ to established ways of doing things. Thus, it is usually argued that firms are likely to shift production to lower cost locations in the later stages of industry evolution. A similar version of the industry-driven cluster cycle emphasises that technological knowledge becomes more standardised and homogenous over time. Variety in both the industry and cluster decline and clusters become prone to overly strong ties and lock-in through time. In some accounts networks also become more coherent and closed so that cluster absorptive capacity also falls (Ter Wal and Boschma, 2007), or the number of firm networks shrinks so the probability of inventing new products and processes is reduced. Thus in this industry-technology approach, when new products emerge they create new clusters that challenge and undermine the old centres (Brezis and Krugman, 1997). An alternative model claims that if new technologies prove convergent with the older ones that dominated declining
clusters, then the latter may be revived by the entry of new firms (Swann, 1998).

A second interpretation of the cluster lifecycle moves away from an industry-driven cycle and suggests that the logic that drives cluster development can be found in more autonomous processes specific to clusters themselves. The simplest version of this approach argues that cluster life-cycles are determined by the balance of agglomerative advantages and disadvantages, and or localisation externalities may themselves have a 'life-cycle' (Potter and Watts, 2008a, 2008b; Neffke, 2009). If there is too much agglomeration, the cluster declines irrespective of the stage of the industry cycle. Some, for example, have identified agglomeration costs even in youthful industries such as biotechnology. Other versions of the independent cluster cycle view are more sophisticated and are based on differences between cluster sustainability within the same industry, so that during a course of an industry’s evolution different clusters rise and fall. Iammarino and McCann (2006) argue that clusters evolve by moving between different types of organizational regime – pure agglomerations, industrial complexes, old and new social networks. In their view movement between these typologies is not determined by the age of the industry, nor by technological and knowledge features alone: “It must be made clear that differences in cluster types and also in cluster evolutionary paths, where they exist, are not necessarily related to industrial sectors” (p. 1032). Menzel and Fornahl (2010) contrast the booming computer sector in Silicon Valley with the shrinking industry in the Boston region in the 1990s and also conclude that, under the same overarching market and technological conditions, some clusters can grow while others fade. They suggest that cluster life-cycles are therefore analytically distinct from industry life-cycles. Again, they also add that clusters can escape the impact of technological decline by diversifying into a new industry, thereby producing a divergence between cluster and industry fortunes, and giving a cluster a new lease of life (in effect launching it on a new ‘life-cycle’):

Thus, there is a cluster inherent component that has an effect on growth and decline of a cluster independent of the particular industry life-cycle. This component is the utilisation of the diverse
competencies that differ between clustered and non-clustered firms (Menzel and Fornahl, 2006, page 9).

Pouder and St John’s (1996) approach offers probably the most well developed explanation of cluster specific competences. They argue that clusters decline largely through cognitive isomorphism and the imitation of close rivals, in effect a place-dependent, cluster-specific ‘technological lock-in’ process. As a result, innovation leadership in an industry shifts over time to non-clustered firms that are not subject to such a process.

The ‘industry-driven’ and the ‘cluster-specific’ cycle views are not exclusive. Some theories attempt to combine elements of both. Brenner (2001) for example suggest that exogenous factors, the demand for products and degree of competition determine a cluster’s critical mass, but that endogenous factors based on local symbiotic interactions are also necessary for a cluster to develop. Nevertheless Brenner’s life-cycle still appears to be primarily based on industrial and technology cycles, as he explains cluster decline in these terms: “The usual reason for a such a change in the exogenous conditions is the emergence of a new technology in connection with the appearance of new products that replace old ones” (ibid, page 19).

While there may be disagreement over whether and to what extent industry-specific or cluster specific processes operate in these alleged ways and are sufficient to produce cluster ‘life-cycles’, the latter notion itself is not unproblematic. In particular, in what sense is the notion of a ‘life-cycle’ an evolutionary construct in the context of an entity such as an industry cluster? The metaphor of the ‘life-cycle’ carries obvious biological connotations of the progression of an entity through the various stages of its ‘life’, from birth to death. Whether the growth, maturity and eventual death of a single organism are evolutionary stages, or are best regarded in a more restricted sense as stages of a pre-defined and ineluctable pattern of development, as discussed above, is an open issue. In the case of the product ‘life-cycle’, these stages refer to the various phases in a product’s position in the market (as measured, for example, by sales or market share). Typically, and importantly, the product (or industry) itself is assumed not to change; it is its competitive
Conceptualising Cluster Evolution

position in the market place that grows, matures (reaches a maximum) and then eventually declines (in the face of competition from newer, better or alternative products or industries). In this sense, the life-cycle notion would seem to describe a form of ‘ontogenetic evolution’, that is the development of a particular entity (or organism) from a set of given and unchanging characteristics (or genes).

This is in contrast to phylogenetic ‘evolution’, which refers to the ongoing evolution of a population of entities (organisms), including changes in the composition of entities making up the population and in the pool of the characteristics (genes) of those component entities (organisms). A business cluster is just such a composite system or population of entities, in this case of firms. As a cluster emerges and develops, so new firms enter and some existing firms may exit: typically, the population of firms making up a cluster continuously changes over time. Further, the characteristics of the firms making up the cluster (their products, technologies, routines, business models, etc) may change over time. As Endler and McLellan (1988) point out, both the addition and subtraction of competing entities - and thus the consequential change in the relative frequency of different entities in a system – and changes in the characteristics of existing entities of a system, are key evolutionary mechanisms. Interestingly, Alfred Marshall alluded to this latter process in his discussion of industrial districts (the intellectual progenitor of Porter’s cluster concept):

Every locality has incidents of its own which affect in various ways the methods of arrangement of every class of business that is carried on in it: and even in the same place and the same trade no two persons pursuing the same aims will adopt the same routes. The tendency to variation is a chief causes of progress; and the abler are the undertakers of any trade the greater will this tendency be (1920, p. 355).

From this phylogenetic view of evolution, there is no inevitability that a composite system such as a cluster will trace out a simple life-cycle type trajectory over time.
Not surprisingly, therefore, more recent applications of the life-cycle in cluster studies have tried to regain some credibility by the focusing on the populations of firms that make up clusters, and recognising that the heterogeneity of firm dynamics within these populations may well mean that the actual development paths of clusters can deviate from the stylised life-cycle model. According to Bergman (2007), for example, the cluster life-cycle concept is merely a ‘discussion template’ that may fit some clusters but which also helps us to understand cluster idiosyncrasy and departures from the ‘norm’ (i.e. the archetypical life-cycle template). Similarly, Popp and Wilson (2007), argue that the interrelationships between technological dynamics, market dynamics, structural dynamics and governance tend to create regularities in the life-cycles of industrial districts, but that these are continually subject to the contingency arising from unexpected events and decision-making under sets of constraints. Yet, such refinements of the cluster life-cycle model are just that: they still retain the ‘ideal’ or ‘normal’ concept of the life-cycle as the basic conception of how clusters evolve. They admit a role for contingency, heterogeneity and agency, but seem to treat such effects more as complicating influences and secondary forces rather than as key drivers shaping cluster development over time. Despite its heuristic advantages and refinements, the life-cycle approach struggles to provide the basis of a convincing general theory of cluster evolution. Indeed explanations of cluster development continue to be hamstrung by recourse to under-explained ‘ageing’ analogies and ‘life-course’ metaphors. The emerging synthesis around cluster life-cycles needs a rethink.

Beyond the Life-Cycle Model: Clusters as Complex Adaptive Systems

But what direction should such a rethinking take? In the rest of this paper we explore the scope and potential, but also the limits, of a complex adaptive system approach. There are strong grounds for arguing that industrial clusters represent complex adaptive systems. As composite entities, clusters exhibit several of the generic features and characteristics that are held to define such systems (for a detailed discussion of those features see Martin and Sunley,
Typically, complex adaptive systems are made up of numerous components with functions and inter-relationships that imbue the system a whole with a particular identity and a degree of connectivity or connectedness. The usual definition of clusters, as “geographic concentrations of interconnected companies, specialised suppliers, service providers, firms in related industries, and associated institutions ... in particular fields that compete but also cooperate” (Porter, 1998, p. 197), would seem to accord closely to this conception. Further, the boundary between a complex adaptive system and its environment is neither fixed nor easy to identify, making operational closure difficult, and the system subject to constant exchange with its environment. This is surely the case with clusters, the boundaries of which are indeed often difficult to delineate. Certainly clusters are highly open systems, subject to constant competition from other clusters in the same industry, and to other ‘exchanges’ with their ‘external environment’.

Yet further still, complex adaptive systems are characterised by non-linear dynamics because of various feedbacks and self-reinforcing interactions amongst components, with the result that they are often characterised by path dependence (Martin and Sunley, 2003; Martin and Sunley, 2007). They are also characterised by emergence and self-organisation: that is, there is a tendency for macro-scale (in our case, cluster-wide) structures and dynamics to emerge spontaneously out of micro-scale behaviours and interactions of system components (the individual firms and institutions that make up a cluster) (see Martin and Sunley, 2011). These macro-scale features can then exert influence (‘downward causation’ or ‘supervenient’ effects) on the micro-scale scale components (Sawyer, 2005). And the process of self-organisation gives complex systems the potential to adapt their structures and dynamics, whether in response to changes in the external environment (e.g. external shocks), or from within either through co-evolutionary mechanisms or in response to ‘self-organised criticality’ (certain thresholds or constraints that arise from the system’s own functioning and structure). Clusters exhibit all of these features and tendencies. Clustering leads to the emergence of cluster-wide macro effects and structures – such as various localisation economies and spillovers, and various institutions and organisations - that serve to reinforce the geographical concentration and competitive advantage of the
Conceptualising Cluster Evolution

individual firms concerned. Because of such interactions and feedbacks, clusters would appear to be quintessential emergent, self-organising and path dependent entities (Belussi and Sedita, 2009). Perhaps most significantly, complex systems can produce multiple possible evolutionary trajectories and unpredictable courses of change. As Folke, (2006, p. 257) states “Theories of complex systems portray systems not as deterministic, predictable and mechanistic but as process-dependent organic ones with feedbacks among multiple scales that allow these systems to self-organize”.

However, thinking about clusters as complex adaptive systems raises a number of issues. In the first place, there is the question of system (cluster) definition and identity and how these relate to the idea of evolution over time. According to Cumming and Collier (2005), discussions of change in complex systems often ignore the question of how to maintain a constant definition of a changing system. There is a need, they suggest, to distinguish between systems that maintain their identity over time and systems that effectively develop into new systems and which may change their identity in the process. They argue that some form of continuity through time would seem to be a central component of system identity, and that identity resides in the continued presence, in space and over time, of key components and key relationships. The key components of a cluster are the particular types of similar and related firms – let us say, computer software firms - and their associated institutions. The key relationships are the various network interactions, interdependencies, inputs, outputs, spillovers, and emergent external economies that connect the co-located firms and associations and which turn them into a functioning system, that is a cluster. Now it is possible, indeed likely, as mentioned above, that the composition of the population of firms making up a cluster will change over time, as some firms fail and go out of business and new ones are created. The new firms may employ more advanced technologies, or produce different variants of the products (such as software) that define the specialism of the cluster. Likewise existing firms may upgrade their techniques and products over time. But provided these changes do not move the firms and the cluster into a different product specialism, say out of computer software into a different sector, say biotechnology, or clean energy technology, the identity of the cluster does not change – it remains a
software cluster. The production techniques used, and the design, quality and range of products produced, by the firms that comprise a cluster may evolve over time even though the identity of the cluster remains the same.

If, on the other hand, extant or new firms in a cluster begin to move into new and different product lines, the identity of the cluster begins to change. Much will depend on the extent to which and direction in which this process occurs. It may be a shift into related product lines that draw on the knowledge base or materials used in the previous specialism of the cluster; or it may a move into quite different unrelated products or technologies. Whether such shifts prove successful will of course depend on a host of factors and conditions, not least of which will be the extent to which the externality effects that typify cluster formation and development are relevant to or adapt sufficiently to support the new firms and new product or technology specialism that these firms are seeking to establish. In either case, a cluster can acquire a new identity, sometimes slowly, at other times quite rapidly. The basic point is that clusters can undergo quite different forms of evolution, some of which need not entail a change in the identity (specialism) of the cluster, and others that involve the replacement of the cluster by another of a different specialism. As is the case with complex systems more generally, clusters can exhibit quite varied forms of evolution.

For this reason, Cumming and Collier (op cit) suggest that it is unlikely that any single ‘correct’ model can fully capture the evolution of all complex systems. Rather, they contend, is it useful to have a set of different types of model available to guide conceptual and empirical work. They identify five such ‘meta-models’ of change (evolution) in complex systems. These are distinguished on the basis of whether the system under study maintains a continuous identity through time; whether alternative stable states (implying some form of self-reinforcing dynamic) are possible in the same physical location; and whether the role of exogenous forces and processes is strong or weak (Table 1). Cumming and Collier do not discuss the life-cycle model as such, but it seems useful to add this to their list, since a complex system could in principle map out a life-cycle type of trajectory over time, and it is useful in any case to characterise this model in terms of the three criteria these authors
use to classify complex-system behaviour. They also consider a sixth type, the ‘evolutionary’ model, but reject this on the grounds that they do not regard it as a whole-system model for ecosystems, societies or economies. As will become apparent later in this paper, we beg to disagree, and believe this model type to be potentially the most useful as a framework for thinking about the evolution of regional and local economic systems, including clusters.

Table 1: Different Meta-Models of Complex Systems

<table>
<thead>
<tr>
<th>Meta-Model (Type of Complex System)</th>
<th>Continuous Identity of System</th>
<th>Alternative Stable State of System Possible</th>
<th>Exogenous Forcing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life-cycle</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Random Walk</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Replacement</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Limitation</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Succession</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Adaptive Cycle</td>
<td>Yes/No</td>
<td>Yes</td>
<td>No/Yes</td>
</tr>
<tr>
<td>Evolutionary</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Adapted from Cumming and Collier (2005)

Of these various types of model of complex system behaviour, Cumming and Collier dismiss the random walk, succession and limitation models as having limited applicability to most real world systems. Under the ‘random walk’ model, a system wanders randomly through a multivariate space: its components and dynamics undergo stochastic changes at irregular intervals of time, and there is no particular regularity in system properties, and no continuity of system identity. It is of little relevance to the study of clusters, which in contrast exhibit self-organisation and self-reinforcing feedbacks, are made up of firms that do not switch randomly between different types of product or specialisation, and which can exhibit a high degree of continuity of identity through time. Of course clusters may possibly arise randomly across space (Porter, 1998, p. 238, himself argues that chance events are often important in the birth of a cluster), though we would contend that few in fact
are the product of purely adventitious or happenstance events, but instead are often the outcome of path-dependent factors and conditions that are themselves place-dependent (see Martin and Sunley, 2006; Martin, 2010). And, of course, a cluster’s firms may be subject to random external market and related shocks. But as functioning local economic systems, clusters are far from random entities.

Similarly, neither the ‘succession’ nor ‘dynamic limitation’ models seem to have much applicability to clusters. Succession models of course have a long history in ecological studies, though they have also been invoked in other disciplines, including economics where they emerge in the idea that economies evolve successively from agricultural to industrial to service (and thence to informational?) forms over time. Succession models imply an endogenously generated progression through a series of steps or stages, or functional forms, often with the assumption that this process leads to some eventual stable form. This does not seem to describe how many clusters evolve. The dynamic limitation model refers to systems the behaviour of which is determined by boundary conditions. Such systems oscillate between phases of growth and contraction as a result of constantly pushing against external limits. Generally, there is no change in form or function – no change in system identity – but a movement back and forth between the limits that act on the system. It may well be the case that clusters encounter limits to their growth and expansion. These could be external or internal to the cluster. For example, the saturation of the market for the products of a cluster’s firms will halt cluster growth and this may well set in motion a selective rationalisation and closure of firms in the cluster, and the setting in motion of an endogenous process of progressive contraction and decline. Alternatively, endogenous limits may emerge that impose constraints on continued expansion of a cluster. Rapid expansion of a cluster can lead to rising inflation of costs, such as wages and land and house prices, and indeed other congestion costs and diseconomies, which then act as a brake on further growth, and may even lead to firms moving out of the cluster to cheaper, less congested locations elsewhere (Stam and Martin, 2011). Or yet again, as a cluster develops there may be progressive lock-in or homogenisation of the knowledge base (through mimetic or convergent behaviour among firms – what Porter (1998, p. 244) calls ‘group think’),
leading to a loss of innovative dynamism and a slowdown in the growth or even stasis of the cluster. It may be possible for the cluster to regain its former growth momentum and even expand again, but it is equally likely that once contraction and decline sets in, the loss of competitive advantage and markets will become cumulative and self-reinforcing, thus making the revival of the cluster back to its former size and strength increasingly more difficult.

One possibility in such instances is that the decline and break-up of a cluster may provide an opportunity for the emergence and development of another in its place. This would be an example of the ‘replacement’ model discussed by Cumming and Collier (op cit). According to this model, complex systems follow after one another in the same location, and may even be similar to one another in certain respects (for example, the new system may retain some of the components of the old), but do not display continuity of identity over time, and therefore constitute different systems. Attention focuses on the constraints or legacies left by the old system and how these condition the nature of any new system that arises to replace the old. Put another way, there may be location-specific path dependence effects involved in the system replacement process. Instances can be found where the collapse and break-up of a particular cluster leaves behind certain resources and competences (such as labour skills, knowledge pools, etc), that have been successfully recombined and adapted as the basis of the development of a new industrial cluster. Indeed, as Martin and Sunley (2006) and Martin (2010) argue, the creation of new industries in particular places is often a path dependent process, arising from the re-use or upgrading of technological, organisational and knowledge assets left by former local industrial paths: new clusters can indeed arise Phoenix-like from the ashes of old, declining ones (Christopherson, 2009).

But the replacement model, like the others discussed thus far, even if applicable, clearly deals with special sorts of cluster evolution or dynamics. Is there a more general or flexible model that admits of a wider range of possibilities, and which can subsume replacement as a special case? The ‘adaptive cycle’ model would seem to offer some scope in this direction, and merits a separate discussion.
The Scope and Limits of the Adaptive Cycle Model

The adaptive cycle model is arguably one of the few well-defined and well-supported interpretations of complex system dynamics. It was initially developed to describe the evolutionary dynamics of ecological systems, but has since been applied more widely in ecological economics and social-ecological studies (Holling, 1986, 1996; Petersen, 2000; Holling and Gunderson, 2002), and is a key element of so-called ‘panarchy’ theory (Gunderson and Holling, 2002). Panarchy is a conceptual framework intended to account for the dual and seemingly contradictory characteristics of all complex systems: stability and change. It has tended to focus primarily on local or regional ecosystems and how these evolve by adapting to the impact of change occurring at a variety of scales: smaller and faster changes at local scale levels, as well as larger, slower, changes at region-wide and supra-regional and global levels. Particular attention focuses on the notion of system ‘resilience’ to external change and shocks, and how resilience itself changes as the system evolves. Resilience is defined in terms of the adaptability of the system. The contention is that complex adaptive systems, such as ecosystems, are characterised by two conflicting tendencies: on the one hand there is tendency in such systems towards increasing internal connectedness and order (or ‘inter-relatedness’) among system components; but, on the other hand, increasing connectedness and order tend to reduce the adaptability (flexibility of response and transformation) of the system to changing environmental or external conditions. This implies that there may be a trade-off or conflict between connectedness and resilience: the more internally connected is a system the more structurally and functionally rigid and less adaptive it is.

The ‘adaptive cycle’ model seeks to reconcile this contradiction or conflict by positing a four-phase process of continual adjustment in ecological, social and environmental systems (Figure 1). Each phase of an adaptive cycle is characterised by varying levels of three dimensions of change: (i) the potential of accumulated resources available to the system; (ii) the internal connectedness of system components; and (iii) resilience, a measure of system vulnerability to and recovery from shocks, disturbances and stresses. These
dimensions or characteristics change with and help define the four phases of evolution of the system (Figure 3). The first stage is the so-called ‘exploitation’ (or ‘$r$’) phase of rapid growth and resource accumulation. This is the phase when an identifiable system emerges and develops. The interconnectedness of system components increases as the system grows, and resilience (adaptability) is high.

![Diagram of adaptive cycle model](image)

**Figure 1: An Adaptive Cycle Model of the Evolution of a Complex System**

This phase eventually merges into the second, so-called ‘conservation’ (or ‘$k$’ phase) of increasing stability, growing stasis and rigidity: the system becomes well established and eventually stabilises around a particular form, structure and mode of self-reproduction. The degree of inter-connectedness or inter-
relatedness is now high, and this reduces or limits the resilience of the system, its ability to adapt to changes or shocks in the external environment. If such a shock should occur, or if stasis and rigidity itself sets off internal processes of decay or atrophy, the system enters a third or ‘release’ (or ‘Q’) phase and declines and contracts. Resources are lost, and connectedness declines. This is then followed by a fourth reorganisation (or ‘α’) phase of reconfiguration, experimentation and restructuring, in which accumulation of resources is slow, connectedness is low, and resilience is increased because several options may be open to as to which direction and in which form a new system develops. In the growth and conservation phases change is gradual and predictable, while in release and reorganization change can be fast and is unpredictable (Peterson, 2000).

Figure 2: Stylised Evolution of Resource Accumulation, Connectedness and Resilience of a Cluster over an Adaptive Cycle

Resource Accumulation: refers to accumulation of productive, knowledge and institutional capital
Connectedness: refers to extent of trade and untraded interdependencies among cluster firms
Resilience: refers to capacity of firms to respond flexibly to shocks internal or external to cluster

Three Scenarios following Release and Decline phase are shown: A – cluster disappears; B cluster undergoes phase of renewal; C- new (different or related) cluster emerges and replaces old

Discussions of the adaptive cycle model in ecological settings admit of two possible outcomes of the reorganisation phase: renewal or replacement. In the
first, the system re-establishes itself, and begins a new cycle of growth and accumulation of resources. In the second, the old system is replaced by a new, with a different identity and function. In essence this a particular instance of the replacement model described earlier, though here it is the outcome of, and part of, an adaptive cycle process. As in the simple replacement model, the new system may incorporate elements and components left over from the old system. If those legacies are substantial, the new system may fall in between renewal and replacement. Finally, resilience, the ability of the system to retain its form and function, is a key feature of the adaptive cycle model, and varies through the cycle, being lowest in the conservation and release phases, and highest in the reorganisation and exploitation phases (Holling and Gunderson, 2002; Pendall, et al, 2008).

The adaptive cycle model appears to have some potentially interesting implications for understanding the evolution of clusters. The growth and development of a cluster is fundamentally about the accumulation of key resources – specialised productive capital, specialised knowledge, and associated specialist supporting institutions. The model suggests that as the accumulation of resources progresses the degree of connectedness of the system components increases (see Figures 1 and 2). It is precisely the development of locally-specific and locally-based interdependencies, both traded and untraded (such as knowledge spillovers and other network externalities) among co-located firms and institutions that is alleged to be the defining feature of a cluster. Yet, the model also suggests that the degree of this interdependence may eventually reach a point where it can undermine the resilience of the firms in the cluster, that is their ability to resist and respond flexibly to external shocks, such as the rise of new competitors or technologies. One can think of various forms this might take. One example might be a cluster composed of a myriad of firms organised into a detailed horizontal inter-firm division of labour, such that it is difficult for one firm to change its particular role, technology or product without corresponding changes occurring across all or most of the other firms. The high-inter-relatedness of the firms comprising the mature textile clusters in North West of England in the early-decades of the twentieth century is often cited as a reason for the failure of those clusters to adjust to intensifying international competition, and
hence of their consequential decline (Frankel, 1920). Another example might be the ‘hub-and-spoke’ type cluster of a dominant firm and its population of local specialist component suppliers. The latter become locked into the hub producer, and this structure too can become characterised by rigidity and inflexibility. These and other types of mature cluster might remain in a stable ‘conservation’ (k) phase of development for a considerable period of time, depending on just how inflexible they become as an economic system and on the nature of any shocks their firms are subjected to.

Since the firms in a cluster will invariably be competing in external (often global) markets, what matters is how they respond and adapt to major competitive shocks. Indeed, the extent of any rigidities and inflexibilities amongst a cluster’s firms often only becomes apparent once the cluster is subjected to an external shock of this kind, and may not be obvious before then. Much depends on the resilience and adaptability of the cluster’s firms. If the resilience of the cluster is low and a major shock occurs, the cluster will move into a ‘release’ phase’, in which firms close, there is disinvestment, little if any growth occurs, and a general contraction in the scale of the cluster takes place. Three outcomes are then possible (see also Trippl and Otto, 2009). The first is where any surviving cluster firms upgrade their products, productivity and competitiveness, regain markets and profitability, and the cluster – though perhaps smaller – succeeds in gaining a ‘second wind’ and growth and accumulation resume (for example, path B in Figure 2). In effect, the cluster adapts through a process of restructuring of its existing firms, and potentially undergoes a new lease of life (i.e. it enters a new α-phase). It may still not succeed, however, and could possibly failure to develop sufficient new growth momentum.

A second possibility is where the old cluster declines and disappears but is replaced by a new one, possibly in a related industrial specialism or one that builds on some of the resources inherited from or left behind from the cluster that has declined (for example, path C in Figure 2). In this case the adaptation is one of renewal around a different type of specialism and the emergence of new firms based on this activity: again a new α-phase is initiated). Thirdly, of course, there are also numerous instances where a cluster goes into decline
and no new local cluster emerges, Phoenix-like, from the ashes (path A in Figure 2). The cluster continues to decline and eventually the industry disappears at that location.

The adaptive cycle metaphor would thus seem to offer several advances on an orthodox life-cycle model. First, it represents a heuristic meta-model of a continuous dynamic process but it does not claim to describe a rigid, predetermined path or trajectory (Holling and Gunderson, 2002). It is recognised that this process (and the interactions involved) will vary significantly between ecological and social and economic fields and in particular that greater reflexivity in human systems will intensify cross-scale interactions. Progress through the four stages of the cycle is not inevitable because cycles may become maladaptive. For example, a system that develops only low connectedness, and fails to accumulate resources and potential (possibly because of ill-advised policy or a difficult institutional context), may have low resilience and become stuck in what is termed a ‘poverty trap’, a pathological departure of the system from the normal adaptive cycle into a state in which it is subject to recurring instability because of lack of connectedness to imbue the system with the capacity to adapt (Gunderson and Holling, 2002; Allison and Hobbs, 2004; Carpenter and Brock, 2008). This can be applied to industrial clusters: a cluster may fail to develop sufficient critical mass and connectedness (localisation externalities and spillovers) to enable innovative adaptation and growth to take place, so that the cluster continually struggles to survive. It is not inevitable, therefore, that clusters invariably follow the adaptive cycle four-stage pattern of evolution, nor that cluster decline is invariably followed either by renewal and rejuvenation of the original cluster, or by its replacement by a new and different cluster (Bergman, 2007).

Second, the adaptive cycle concept suggests that change is not imminent and predetermined but instead is unpredictable as a result of interactions between nested or embedded systems which act to reshuffle system configurations and provide novel experiments. For instance, the consequences of a collapse of a component subsystem might be expected to vary according to the phase reached by the system as a whole. If the wider system is in a period (phase) of
low resilience then a micro-scale ‘revolt’ may cascade upwards and overwhelm the regularities operating at higher scales, across the system as a whole. This clearly applies to clusters; if an important ‘hub’ firm within a cluster collapses or reorganizes, or shifts production to another site, an internal shock of this kind could either lead to a wider collapse of a vulnerable cluster, or alternatively to a reorganisation of resources and greater opportunity for surviving firms if the cluster concerned is in a growth phase.

Third, the adaptive cycle perspective foregrounds the importance of recombination and reuse of resources. According to the metaphor, the unpredictable recombinations and linkages that emerge during reorganisation phases are tested during the longer periods of conservation. Renewal depends on the reworking of the legacies from preceding cycles. This also allows opportunities for the incorporation of exotic and novel elements. This has direct relevance to the study of clusters, of course, as many have argued that the origins of new clusters can partly be traced to the reuse and recombination, or ‘rebundling’ of legacies of the past and the potential of older clusters (e.g. Bathelt and Boggs, 2003; Boschma and Wenting, 2007). Elsewhere, indeed, we have argued that precisely because of these legacies there is a degree of both path and place dependence in the creation of new cluster trajectories (Martin and Sunley, 2006). For example, the outdoor clothing and equipment specialisation that has grown in the North West of England has incorporated legacies from both the preceding Lancashire textiles cluster and metalworking districts elsewhere in the North (Parsons and Rose, 2005). Without some inherited potential the opportunities for the creation of a new cycle of cluster growth are significantly constrained.

But while undoubtedly suggestive, the adaptive cycle model is not unproblematic, however, both as a model of how complex systems evolve in general, and of cluster evolution more specifically. At a general level, Cumming and Collier (2005) point out that the model does not fit all types of ecological systems, and that it leaves much unsaid about detailed processes. Does the metaphor primarily describe a regular sequence of change post hoc or does it actually explain the mechanisms causing this sequence? Moreover, there are a number of ambiguities about the adaptive cycle metaphor which
mean that it is arguable whether it captures the key features of complex adaptive systems. A key defining signature of complex adaptive systems is a phenomenology that is characterised by a multiplicity of possible outcomes, endowing such systems with the capacity to choose, explore and adapt. According to Nicolis and Nicolis (2008, 2009), this process manifests itself in two ways. One is the emergence of traits encompassing the system as a whole that in no way can be reduced to the properties of the constituent components. In our case, a cluster is more than the sum of its parts, as Porter (1998) himself emphasises. Another is the intertwining, within the same phenomenon, of large-scale regularities and seemingly erratic evolutionary trends. This co-existence of order and disorder significantly restricts the predictability of the evolution of the system. And the greater the heterogeneity of firms making up a cluster, then the less predictable that cluster’s evolutionary path might be expected to be. In addition, the evolutionary dynamics of complex systems, including clusters, are shaped by the interplay of ‘downward’ and ‘upward’ causation (de Haan, 2007). As a result of this continual two-way causation, the evolutionary path taken by a system can undergo frequent disruption and redirection: hence again, its unpredictability.

The adaptive cycle approach undoubtedly reflects some aspects of complex systems. It explicitly rejects a single, cumulative and imminent logic determining system evolution and instead emphasises that system change is an outcome of the balance between experimentation and novelty, and conservation and selection, as well as the interactions between entities at different scales. Yet at the same time, the approach insists that system change is controlled by a handful of key variables (Holling, Gunderson and Ludwig, 2002). In a similar way, most life-cycle approaches to clusters search for key controlling variables in cluster evolution. Recent work on clusters has seemed reluctant to let go of the idea that there is a dominant dynamic that determines the evolution of clusters. Ter Wal and Boschma (2009), for instance, argue that if we are to understand cluster evolution we have to pay careful attention to the heterogeneity of firms within clusters and unravel the complex co-evolutions of firms, networks and industries. We would endorse that view. However, they conclude that the phasing of these co-evolutions is governed by the industry life-cycle in which the cluster is embedded. Firms, networks and
industry structure all appear to evolve together with some synchronicity as the industry matures and networks gradually close. Several studies likewise portray the knowledge variety of a cluster as its dominant dynamic. Here mimetic behaviour amongst firms causes genuine knowledge variety to decline as a cluster matures, the argument being that firms learn most from their neighbouring firms, and as a result technological lock-in occurs, with negative consequences for radical, growth-enhancing innovation. Life-cycle approaches to clusters have tended to envisage a sequence in downward causation effects, usually from cluster to firm. Thus being co-located in a cluster first positively affects firms and then negatively constrains them. In Menzel and Fornahl’s (2009) model, for example, clustering initially boosts absorptive capacity but then later becomes a negative influence as firms learn most from their neighbours and their heterogeneity declines. Again, while it is undeniable that such effects are evident in certain clusters, we are sceptical that downward causation will generally move in such a predictable way in complex systems where it operates in a shifting tension with processes of upward causation (such as from firm to cluster). Thus, if neighbouring firms have built successful open innovation networks there is no reason why knowledge heterogeneity should decline. As Tappi (2005) shows, ‘extrovert’ entrepreneurs who have developed a broader cognitive framework by participating in non-local networks may rejuvenate a cluster. Thus the Marche music cluster moved from accordion production into electrical instruments and then diversified into other electronic products. Even in mature clusters firms may produce innovations through mistakes in imitation and unintended outcomes (Staber, 2007) and, on occasions, such innovations have important effects on the viability of their cluster. The problem, then, with assigning too much dominance to one or two key controlling factors is that disruptive processes of upward causation are almost completely ignored.

In contrast, the adaptive cycle tries to capture both capture and downward causation but it appears to confine upward causation to two key episodes. For most of the adaptive cycle, hierarchical structures regulated by a small set of critical processes and variables are in control, as larger and slower cycles are pictured as constraining the behaviour of smaller and faster cycles. Although large scale collapses can cascade downwards and undermine the successful
adaptation of smaller systems, for the most part larger scale systems are typically seen as conservative and stabilizing forces which act to conserve the memory of successful surviving experiments. By supplying resources and potential larger scale systems help smaller cycles to ‘remember’ and restore their growth during their reorganisation phases. As Holling, Gunderson and Ludwig (2002, p. 20) write, “In essence, larger and slower components of the hierarchy provide the memory of the past and of the distant to allow recovery of the smaller, faster adaptive cycles”. However, this pattern of control and restoration by slower and larger systems is disrupted during episodes of release and reorganization. In these two rapid transitions, larger scale systems are highly sensitive to small disturbances and experiments which provide brief windows for experimentation and the generation of novelty. In the words of Holling and Gunderson (2002, pp. 72-73) “During other times, the processes are stable and robust, constraining the lower levels and immune to the buzz of noise from small and faster processes. It is at two phase transitions between gradual and rapid change and vice versa that the large and slow entities become sensitive to change from the small and fast ones”. In the release and the reorganization phases smaller scale systems can cause a ‘revolt’ and small changes can cascade upwards with long term consequences.

The question raised is whether cluster evolution shows the same episodic nature and whether upward causation occurs mainly in these faster critical phases. Do the effects of firm decisions, experiments and reorganizations cascade upwards to the rest of a cluster primarily during these phase transitions? Are cluster evolutions particularly sensitive to firm decisions during a period of release and renewal? Certainly one might reasonably argue that the decisions of individual entrepreneurs and firms may have either strong negative or positive feedbacks during these episodes. For example, a key innovation by a particular firm may have large consequences during a period of reorganization as it might trigger a cumulative process towards a new technological trajectory, or a new branch of an existing trajectory. For example, there are numerous established industrial sectors, such as computers, pharmaceuticals, medical devices, machine tools, cars, cameras, even banking, that have been revolutionised and renewed by the creation of new families of products (Nelson, 1996; Windrum, 2005). Such sectors have
shown repeated rounds of product diversification and reinvention through the exploitation of product niches. In such cases we might well find that innovations by clustered firms have reshaped market demand and disrupted and complicated the trajectory of the entire industry in question. However, it is unknown whether this type of renewal can be traced to an innovation during a particular stage of cluster evolution, or whether such renewals and experiments emerge continually and repeatedly throughout the course of a cluster’s development.

It is also not clear whether the idea of panarchy, composed of nested adaptive cycles, provides a useful way of thinking about the co-evolution between a cluster and economic and institutional systems at broader scales. The notion of periodic interactions between a hierarchy of lumpy systems appears to rather constrain our appreciation of the complex interactions between a cluster and its environment. In the case of clusters, not only does the external competitive environment impact on a cluster’s firms, the activities of those firms can impact back on that competitive environment across the industry more generally. That external environment consists of the (global) markets for the products or services provided by a cluster’s firms, the various policy regimes (both national and global) that comprise the regulatory environment (such as standards, trade restrictions, etc) within which the cluster industry operates, and of course competitors (and collaborators) in other clusters in the same or related industry (Figure 4): often such competitors or collaborators located overseas. For example, a given footwear, or biotech, or computer software cluster, is in competition (and possibly also collaboration) with other such clusters across the globe: it is affected by the actions of those other firms (the external environment), but at the same time is part of the eternal environment faced by each of those other firms and clusters. Indeed, a key interaction between a cluster’s firms and other similar or related firms in other clusters elsewhere concerns knowledge flows. It is normally claimed that local knowledge spillovers amongst a cluster’s firms are a defining feature of a cluster - the idea of the cluster as a ‘knowledge community’. But it is increasingly recognised that while local spillovers may exist, external sources of knowledge from competitors or collaborators far removed from the cluster in question may be just, if not more, important drivers of innovation in a
cluster's firms (Bathelt, Malmberg and Maskell, 2004). And of course, this process operates in the reverse direction, outwards from the cluster in question to external competitors and collaborators. This two-way interpenetration of, and in a sense blurring of the boundaries between, a

![Diagram of Two-way Interactions between a Cluster and its External Environment]

**Figure 4: The Two-way Interactions between a Cluster and its External Environment**

system and its environment is a characteristic feature of complex adaptive systems. Indeed it is recognised as central to the process of evolution itself:

> At every moment natural selection is operating to change the genetic composition of populations in response to the momentary environment, but as that composition changes it forces a concomitant change in the environment itself. Thus organisms and environments are both causes and effects in a co-evolutionary process. (Lewontin, 2000, p.126)
Yet most discussions of the adaptive cycle model treat systems as if they can be neatly allocated to distinct spatial and temporal scales, and tend to ignore or downplay the impact of that system’s dynamics and development on its external environment. Moreover, given the reflexivity and non-linear nature of economic systems at large scales, as illustrated so vividly in the recent global financial crash, there are no grounds for assuming that larger scale economic systems typically operate at slower speeds than smaller ones.

The episodic nature of the adaptive cycle model is further demonstrated by the assumption is that it is the occurrence of an external shock during the ‘conservation’ phase, when the system’s resilience (adaptability) is low, that de-stabilises it and moves it into the next, ‘release’ phase of the cycle. The model has much less to say about other possibilities that can occur in the case of industrial clusters. One such possibility, which we have referred to above, is where there are endogenous emergent mechanisms inherent in the very growth of the cluster – the onset of negative externalities, or lock-in to a particular dominant production technology that ceases to yield increasing returns, or the growth of labour power which holds back productivity enhancing improvements in production, or some other such effect that arises through the high degree of connectedness of cluster firms – which become so strong and prevalent that they cause the cluster to lose its growth momentum. This possibility is akin to the idea of ‘self-organised criticality’ found in the theory of complex systems. In other words, a major external shock may not be needed to destabilise the cluster and cause the onset of its decline. That is to say, irrespective of conditions in the external market for which a cluster’s firms produce, those firms may simply lose their dynamic because of the build up of negative local externalities or because of the exhaustion of productivity-raising improvements. Recent research on the much-celebrated Cambridge (UK) high-tech cluster has identified what appears to be the emergence of such endogenous constraints and barriers to its further growth, including acute road congestion problems, high land costs, and a lack of a sufficiently growth-orientated management culture (Simmie, Martin, Wood and others, 2006; Stam and Martin, 2011). In addition, we can also envisage situations where the external shocks to a cluster are more or less continuous: competitors are constantly challenging the market position and hence profitability of a
cluster's firms; or the cluster's firms may be in a sector where ongoing technological change is a defining feature of that sector, and not some occasional or periodic event.

Towards a Modified Adaptive Cycle Model of Cluster Evolution

We have argued then, that the adaptive cycle when applied to cluster evolution raises some important questions about the interactions of systems at different scales as well as about the re-use and re-absorption of resources from earlier phases of growth, and highlights the need to understand the complexity of interactions between upward and downward causation. The metaphor strongly suggests that we need to focus on what happens during phases of ‘release’ and ‘renewal’, possibly in order to examine whether upward causation and firm decisions are especially significant and consequential during these phases. At the same time however, the adaptive cycle model is rather ‘lumpy’ as an evolutionary framework in that it distinguishes discrete systems at different spatial and temporal scales and then proposes that their interactions are inherently episodic. At present we do not know how much of this episodic ontology can be applied to the evolution of clusters. The complex interactions witnessed between clusters and their environments mean that the idea of a single underlying sequence of expected cluster development remains too predictable and, instead, we should recognise that there are numerous development trajectories whose realisation depends on contingent and strategic decision-making by cluster-based firms. All potentially imply different evolutionary dynamics of cluster resource accumulation, connectedness and resilience. In what follows we therefore argue that the adaptive cycle model as applied to clusters should be modified in order to emphasise a number of different possible sequential trajectories.

The modified and extended model shown in Figure 3 is designed to retain some of the useful stylised regularities highlighted by the basic adaptive cycle model whilst at the same time allowing greater scope for the critical influence of contingent, agency, and firm response to pressures. The modified model
admits of six possible alternative evolutionary trajectories (see Table 2). In the first place, of course, the classic ecological cycle of stabilisation, decline and release, re-organisation and growth (\(\alpha-r-k-\Omega\)), clearly should be retained as one possible sequence as it seems to fit some observed cases of cluster evolution. For example, it has been shown that the decline of the tyre cluster around Akron, Ohio, has had a productive legacy in the sense private and public research facilities have moved from rubber into polymers (Carlsson, 2001). In this sense resources released by the demise of tyre making were to some degree reorganised into a new specialisation. The growth of low carbon technology industries in the Ruhr, Germany, is similar in some respects. On a smaller scale, the shrinkage of the Birmingham jewellery quarter has, it is claimed, also released resources that have become the basis of a creative district focused more on specialist retailing (De Propris and Lazzeretti, 2008). In similar fashion, a subset of the resources released by the decline of textile and steel industries in Northern England have been shifted into an outdoor equipment and clothing industry (Parsons and Rose, 2005).

**Figure 3: A Modified Cluster Adaptive Cycle Model**
Table 2: Some Alternative Cluster Evolutionary Trajectories under the Adaptive Cycle Model (see Figure 2)

<table>
<thead>
<tr>
<th>Evolutionary Trajectory</th>
<th>Phases of Evolution and Typical Characteristics</th>
<th>Possible Mechanisms</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cluster full adaptive cycle</td>
<td>Emergence, growth, maturation, decline, and eventual replacement by new cluster. Follows the archetypal adaptive cycle. Replacement cluster likely to draw upon resources and capabilities inherited from old cluster.</td>
<td>Resilience rises and then falls as cluster passes through phases of cycle. Cluster atrophies either because of internal rigidities or exhaustion of increasing returns effects, or is unable to withstand major external competitive shock. But sufficient resources, inherited capabilities and competencies are left to provide basis for emergence of new cluster based on related or cognate specialism.</td>
</tr>
<tr>
<td>2. Constant cluster mutation</td>
<td>Emergence, growth and constant structural and technological change. Cluster continually adapts and evolves, possibly by successive development of new branches of related activity. Particularly likely where basic technology has generic or general purpose characteristics.</td>
<td>Cluster firms are able to innovate more or less continuously and cluster constantly mutates or widens in terms of industrial specialisation and technological regime. High rates of spin-offs from existing firms and spin-outs from local research institutes or Universities. Cluster has high degree of resilience.</td>
</tr>
<tr>
<td>3. Cluster stabilisation</td>
<td>Emergence, growth and maturation, followed by stabilisation, though possibly in a much reduced and restricted form. Cluster might remain in this state for extended period of time.</td>
<td>Though cluster possibly experiences a phase of decline in scale, remaining firms survive by upgrading products and/or focusing on niche or prestige market segments. Cluster retains a modest degree of resilience, but remains potentially vulnerable to (further) decline.</td>
</tr>
<tr>
<td>4. Cluster re-orientation</td>
<td>Upon reaching or nearing maturation, or upon onset of early cluster decline, firms re-orientate their industrial and technological specialisms, and new cluster emerges.</td>
<td>Cluster in effect branches into new form without going through long period of decline. The more innovative and externally connected lead firms may play key role in this process, for example reacting to market saturation or rise of major competitors, or a technological breakthrough may activate re-orientation.</td>
</tr>
<tr>
<td>5. Cluster failure</td>
<td>Emergent cluster fails to take off and grow. Any remaining firms do not constitute a functioning cluster.</td>
<td>Cluster fails to achieve sufficient critical mass, externalities, or market share. Strategies of 'anchor firms' may weaken the cluster and innovation may also falter. New firm formation low and/or firm failure rate high, which deters new entrants.</td>
</tr>
<tr>
<td>6. Cluster disappearance</td>
<td>Emergence, growth, maturation, decline and elimination. No conversion into or replacement by a new cluster. Classic life-cycle trajectory.</td>
<td>Cluster experiences same eventual atrophy and decline as in the full adaptive cycle pattern (1. Above), but inherited resources and competences not sufficient or ill-suited to form basis of new cluster formation leading to a deep 'poverty trap' and disappearance.</td>
</tr>
</tbody>
</table>
But in addition to this classic sequence, the modified model suggests there are also other widely experienced trajectories of cluster change, and that specific clusters may actually have highly differentiated histories as they combine these cycles in different ways and at different speeds. One of these stylised alternative evolutionary courses is indicated by the circle labelled *constant mutation* in which cluster firms more or less continually develop new products or product variants through on-going innovation, and the cluster does not stabilise into a mature pattern of strong networks and accumulated resources around a particular specialism, but progressively evolves – in an adaptive path dependent manner (Martin, 2010). In this case, the firms in an established cluster may prove flexible enough to respond and adapt to a constantly changing market and technological environment, the cluster may remain in the ‘exploitation and growth’ $r$-phase of the adaptive cycle model (that is, the evolutionary path is $\alpha-r-r'-r''\ldots$), and need not encounter the rigidity and stasis that would indicate the onset of a ‘conservation’ phase. Of course, not all clusters will be able to be so flexible and adaptive. Much will depend on the nature of the industrial specialism on which the cluster is based: high-tech clusters, for example, might be expected to be far more flexible and adaptable than clusters specialising in sectors or activities in which there are large fixed (sunk) capital costs, or those in which there are strict limits to how far the threat posed by cheap labour competitors elsewhere can be offset by productivity enhancing technological improvements in production (for example, Schamp, 2005). Indeed it has been argued that the new ‘open networks’ of high technology clusters, especially in life-sciences and computing and ICT, may enable very rapid adaptability and collective learning. Certainly the most successful of these open network clusters, such as Silicon Valley and Medicon Valley (see Moodysson et al, 2008) have been able to exploit new knowledges and generate repeated waves of new products. A possibly important mechanism here may be that such networks allow technological ‘pre-adaptation’ in which firms accumulate a diversity of technological knowledge, some of which at a later date proves important to cluster mutation and rejuvenation (Baglieri, Cinici and Mangematin, 2010). Whether such ‘exploitation circles’ are sustainable over decades, however, is a key question for future research. For example, the Cambridge, UK, high-technology cluster has in the past been seen as one of these youthful constantly
mutating clusters (Garnsey and Heffernan, 2005) but, as noted above, recent work has started to identify signs of constraint, slowdown and rigidity (Stam and Martin, 2011).

Another alternative trajectory is that of cluster stabilisation. There are a number of cases where the cluster eventually decreases in size but nevertheless does not rigidify and disappear. In effect the cluster stabilises and remains in the $k$-phase of the adaptive cycle model (so the evolutionary path is $\alpha-r-k-k'-k''$ ....). For example, Bryson, Taylor and Cooper (2008) explain how the lock-manufacturing cluster in the West Midlands has survived by switching market focus to concentrate on high value-added niche markets and just-in-time delivery to specialist buyers. Similarly, it has been argued that some Italian districts have stabilised by moving from the production of final goods to the production of machinery, while others have upgraded the quality of goods and specialized to a greater degree in design and marketing (Rabellotti, Carabelli and Hirsch, 2009). In some cases of cluster stabilisation, constituent firms have diversified but not to such a degree that they have rejuvenated or launched related trajectories. For example, Chapman, MacKinnon and Cumbers (2004) describe how small and medium enterprises in the Aberdeen oil complex have diversified into export markets but on the whole have not moved into radically new markets. Many clusters show only incremental innovation and seem precariously close to decline, but nevertheless manage to survive in a reduced form for decades.

A third stylised and contrasting path identified by this modified cycle model is cluster re-orientation in which a mature cluster moves from a stable period to a new phase of emergence and renewal and in effect, short-circuiting the process of decline and release. In such a case, the firms making up a mature cluster would retain sufficient resilience and flexibility to be able to anticipate growing pressures arising from external competition or exhaustion of markets. It may be that the firms begin restructuring and re-orientating their activities before the process of decline gets underway, and in so doing succeed in moving the cluster directly to a new $\alpha$-phase of potential redevelopment (giving an evolutionary sequence of $\alpha-r-k-\alpha$, see Figure 3). For example, the Montebelluna sportswear cluster has been widely discussed as an example of a
district that has been stabilised by radical product diversification that moved the cluster from a traditional focus on boots to technical sports shoes. In this case the attraction of foreign investment has reinforced and confirmed this new cycle of growth and exploitation (Sammarra and Belussi, 2006). In other cases a cluster may undergo a turnaround in response to a perceived start of a crisis. The notion of turnaround implies that firm reactions during the onset of a crisis are crucial, and the effectiveness of their response determines whether decline is averted or confirmed (Chowdhury, 2002). The Boston high-technology cluster represents an example of such reorientation as after the onset of periods of decline it has switched into new specialisms and displayed a track record of turnarounds (see Bathelt, 2001). A further obvious example is the financial services cluster in the City of London which has suffered episodes of decline but has to date managed to re-orient its activities in order to regain profitability.

Finally, our revised adaptive cycle highlights that there are two phases in cluster trajectories when they are vulnerable to terminal decline. The first is the outcome of rigidity, high internal connectivity and ‘lock-in’ leading to non-regenerative decline and cluster disappearance (that is α-r-k-Ω-d). As we have noted already and as the economic history of modern Britain demonstrates only too well, despite the release of resources involved in a cluster’s decline there is nothing inevitable about re-organization into a new emergent pattern and the industrial district may just progressively decline. There are numerous examples including Sheffield steel (Potter and Watts, 2010) Dundee jute (MacKay et al, 2006) and Como Silk (Alberti, 2006) where declining clusters have failed to generate any offspring, or innovate or resorganise themselves out of decline. Typically in such cases, after rationalization the clusters come to be dominated by a handful of large companies whose decisions determine the fate of the cluster, as is exemplified by the Staffordshire pottery and ceramics district (Sacchetti and Tomlinson, 2009).

But such ‘ageing’ is not the only type of cluster demise. In addition there is considerable evidence that clusters are highly liable to fail to become established during the emergent phase when selection pressures are strong.
many cases emergent clusters that have been driven mainly by policy initiatives suffer problems of failing to attract sufficient private investment and lack of social networks among entrepreneurs. The attempt to build a digital cluster in Dublin appears to be a good example (Bayliss, 2007). In other cases the decisions taken by large anchor firms have been crucial to the fate of emergent clusters (Niosi and Zhegu, 2010). Feldman (2003) and Feldman and Lendel (2010), for instance, argue that the growth of biotechnology and optics clusters in the USA has been strongly shaped by the degree to which ‘anchor tenant’ firms have produced entrepreneurial spin-offs and co-evolved with research institutions in order to transform scientific knowledge into commercial products. \textit{Emergent cluster failure} (α-f) then is the final type of stylised cluster trajectories highlighted by our modified adaptive cycle model. What the model does not tell us, of course, is how prevalent these six different trajectories are, or why particular clusters evolve along these different possible trajectories. And clearly, the model does to exhaust the full range of hybrids and combinations that no doubt exist. Further, and most crucially, of course, the different stylised alternatives depicted in Figure 3 depend on the time horizon over which our analysis is made. For example, particular clusters may remain in the constant mutation and rejuvenation phase or the mature stabilisation phase for considerable periods of time; but eventually, they may succumb to the \textit{phase.} Our interpretations and descriptions of cluster evolution are perforce shaped by the time period over which we observe them: do all clusters ultimately disappear? What our modified adaptive cycle cluster evolution model does at least do is suggest that future research needs to carefully examine how the complex co-evolution of firms and their cluster contexts may give rise to a variety of evolutionary paths that are perhaps not best captured by the standard life-cycle model.

Some Conclusions: Reinventing the Cycle or a Possible Research Agenda?

Our argument in this paper has been that the life-cycle model of cluster evolution, though having obvious appeal, may not be the most insightful interpretative framework. While there are no doubt specific processes in
clusters that develop through time, it is not clear that understanding them in terms of a simple life-cycle is very helpful. We do not wish to deny that in some circumstances a dominant mechanism or tendency may govern the viability of a particular cluster, and its evolutionary history may well trace out a conventional life-cycle trajectory. However, we are uncomfortable with the assumption that cluster evolution is universally or even typically determined by the progress and state of one dominant dynamic or process. If clusters are understood as complex adaptive systems in which agents learn through interactions then their emergent dynamics cannot be readily reduced to, nor explained by, one set of components or one dimension of variability. Clusters as complex adaptive systems show a type of reflexive emergence in which agents respond to how they perceive and interpret the system as a whole, and the external environment with which they individually and the cluster as a whole interacts (see Martin and Sunley, 2011b). Hence we need to avoid both epistemological reductionism in which concepts and theories or clusters are reduced to concepts and theories at other levels (such as the industry lifecycle), and methodological reductionism in which cluster systems are seen as determined simply by the properties of their parts (such as local knowledge networks).

Having said that, complex systems models of cluster evolution are as yet very undeveloped. We have briefly reviewed a number of such models, and focused particularly on the adaptive cycle model that seems to have attracted increasing attention recently in ecology and ecological economics. This model has the advantage that it assigns particular importance to the resilience (that is, adaptability) of a complex system, and how resilience both shapes and is shaped by the development of the system concerned. This seems to us to be a valuable idea to explore with respect to cluster evolution. The model also allows for the possibility of system (cluster) renewal (recovery) as well as replacement, or maladaptive collapse into a ‘poverty trap’. Furthermore, the model relies not on the abduction of a biological metaphor, but instead on an ecosystem analogy, which arguably is more appropriate.

However, in its current formulation, the adaptive cycle model also has its ambiguities and limitations. The assumption that the evolution of a complex
system always occurs through a four-phase sequence is somewhat restrictive, and in this respect is open to similar criticisms that can be levelled at the lifecycle model. And while the adaptive cycle model supposedly allows for endogenous as well as exogenous ‘forcing’ mechanisms to move a system through these four phases, most of the emphasis is on the role of an exogenous shock to fulfil this function. Certainly, clusters are subject to occasional major shocks. But most are also subject to more or less continuous external pressures and challenges from market shifts, technological advances and the rise of new competitors and regulatory developments, as well as by endogenous constraints and opportunities. Further, the adaptive model makes only a rather restrictive allowance for the two-way nature of the interaction between a cluster and its external environment. Cluster evolution has to be seen not simply in terms of the development of the cluster in isolation, but in the context of its co-evolution with the (global) industry of which it is itself a part, and other similar clusters elsewhere with which it is in competition. Further, given its origins in the ecological literature, the model is fundamentally episodic and implies that the micro-behaviours – the agency of individual system components (in our case, individuals and firms), are only highly significant for the course of cluster evolution during specific periods of rapid change and transition.

Yet these limitations need not be insurmountable. As Cumming and Collier (op cit) point out, the adaptive cycle model is a ‘meta-model’, a generalised interpretative schema, a way of thinking about change in complex systems. It is not intended as a detailed description or depiction of the precise or specific micro and macro processes involved in a particular real-world system. The implication is that if the model is to be useful for conceptualising the evolution of clusters, it will need to be adjusted and revised to make it more appropriate for this and other types of local economic system. Our revised adaptive cycle model outlines some possible ways in which this adjustment could be achieved in order to recognize a wider and more varied set of stylised cluster evolutionary pathways. It emphasises that the relationships between resilience, connectedness and capital accumulation in clusters are more complicated and unpredictable than the classic adaptive cycle would suggest, and thereby highlights the need for further research into how clusters change.
Just as with the life-cycle model, the adaptive cycle model need not fit all clusters: indeed, as we have stressed throughout this paper, we are sceptical that any such universal model exists. What does seem worth pursuing, however, is the conceptual and empirical exploration of the applicability of a modified adaptive cycle model as part of the research agenda on understanding cluster evolution. We have suggested some possible directions which this exploration might take.

References


