The Place of Path Dependence in an Evolutionary Perspective on the Economic Landscape

Ron Martin* and Peter Sunley**

* Professor of Economic Geography, Department of Geography, University of Cambridge, Cambridge CB2 3EN, UK
E-mail: rlm1@cam.ac.uk

** Professor of Human Geography, Department of Geography, University of Southampton, Southampton, SO17 1BJ, UK
E-mail: P.J.Sunley@soton.ac.uk

Chapter in:
Boschma, R. and Martin. R. (Eds) 
*Handbook of Evolutionary Economic Geography*, Chichester: Edward Elgar


Acknowledgement: Ron Martin’s research for this paper was undertaken whilst holding a Leverhulme Major Research Fellowship.
1. **Introduction: Creating Space for History in Economic Geography - The New Focus on Path Dependence**

Since the notion of path dependence entered the economics lexicon in the 1980s and early-1990s, particularly through the work of Paul David on the economic history of technology (David, 1985, 1986, 1988, 1992, 1993a, 1993b, 1994), and that of Brian Arthur on nonlinear, self-reinforcing economic processes (Arthur, 1988, 1989, 1994a, 1994b, 1994c, 1994d), it has assumed prominence as an evolutionary concept not only in economics itself, but across a wide range of social, organisational, technological and managerial sciences. Some even see the concept of path dependence as a major building block of a new interpretative or epistemological paradigm.

Economic geography has also been swept up in this wave of ‘path dependence thinking’. A number of leading theorists in the subject have argued that path dependence is one of the fundamental features of the economic landscape. For example, according to Richard Walker:

> One of the most exciting ideas in contemporary economic geography is that industrial history is literally embodied in the present. That is, choices made in the past – technologies embodied in machinery and product design, firm assets gained as patents or specific competencies, or labour skills acquired through learning – influence subsequent choices of method, designs, and practices. This is usually called ‘path dependence’... It does not mean a rigid sequence determined by technology and the past, but a road map in which an established direction leads more easily one way than another – and wholesale reversals are difficult. This logic applies to industrial locations as well... (2000, p. 126).

---

1 In economics, ideas akin to ‘path dependence’ can in fact be traced back more than a century, to Carl Menger’s (1883) analysis of ‘institutional emergence’ and Thorstein Veblen’s (1898) concept of ‘cumulative causation’ in the evolution of habits and conventions. And the closely related concept of ‘hysteresis’ has also been around since the mid-1970s (see Elster, 1976; Franz, 1990; Cross, 1993; Katzner, 1993). Outside of economics we find the concept being applied to topics as diverse as decision-making and social behaviour (Anderlini and Ianni, 1996, Goldstone, 1998); industrial organisation (Antonelli, 1997); power generation technologies (Cowan, 1990); pest control programmes (Cowan and Gunby, 1997); industrial technology strategies (Ruttan, 1997; Araujo and Harrison, 2002); technological leadership (Redding, 2002); corporate governance (Bechchuk and Roe, 1999); legal systems and social institutions (North, 1990); historical sociology (Goldstone, 1998; Mahoney, 2000); corporate organisation (Sydow, Schreyögg and Koch, 2005); and politics and state intervention (Bridges, 2000; Pierson, 2000; Dimitakopoulos, 2001; Magnusson, 2001). This proliferation of path dependence ideas can itself be interpreted as a reflection of what some have referred to as an ‘historical turn’ in the social and cognate sciences (Abbott, 2001; Mahoney and Rueschemeyer, 2003; Howlett and Rayner, 2006).

2 This logic had in fact been anticipated somewhat earlier by Storper and Walker who argued that “Localized technological change in an industry can be understood, like all industrial development, as an evolutionary path in which each step moves one way from a past that cannot be recovered and that limits future directions” (1989, p. 113).
Allen Scott (2006) is even more emphatic, and argues that any attempt to understand the economic landscape

must formulate the problem by reference to a dynamic of cumulative causation whose logic is definable not in terms of some *primum mobile* or first cause, but in terms of its own historical momentum. This... points... to the importance of an ontology of regional growth and development that is rooted in the idea of path dependent economic evolution and recursive interaction (p. 85).

Alongside these and other theoretical invocations, path dependence ideas and phraseology have found their way into a growing body of empirical work in economic geography. Thus the idea has been deployed in discussions of the persistence of regional disparities in economic development; of the ‘lock-in’ of regions to particular economic specialisations; the revival and reinvention of former local industrial configurations; and of the emergence and self-reinforcing growth of ‘high-tech’ clusters (see, for example, Grabher, 1993; Storper, 1995; Storper, 1997; Cooke and Morgan, 1998; Bode, 2001; Kenney and von Burg, 2001; Boschma, 2004, 2005; Bathelt and Boggs, 2003; Fuchs and Shapira, 2005; Gertler, 2005; Hassink, 2005a and 2005b). Such economic-geographic studies employing the concept of path dependence reflect a growing interest in the historical dynamics of economic landscapes, a realisation that to understand geographically uneven development, in all of its manifestations, it is necessary to create a space for history.

In one sense, of course, this recognition of the importance of history by economic geographers is not entirely new. The notion of ‘cumulative causation’, which is closely related to path dependence ideas, enjoyed some degree of popularity within the discipline in the 1970s, though unfortunately it has since largely slipped from visibility. Much Marxist economic geography in the 1980s was concerned to explain uneven regional development as an historical process. For example, Massey’s (1984) important work on the spatial divisions of labour was founded on the argument that the economic landscape inherits the legacies of its past development and that these legacies exert an influence on its present and future development. And David Harvey’s central aim was (and still is) to explain uneven regional development as an historical process driven by capitalism’s episodic phases of accumulation and crisis, as a dialectic between preserving the values of past commitments made at a particular place and time, and devaluing them to open up fresh room for accumulation at some future point.
The Place of Path Dependence

in time (Harvey, 1982, 1985, 2006; see also Smith, 1982). But the recent ‘evolutionary turn’ in economic geography is distinctive in that it draws its inspiration explicitly from evolutionary ideas and concepts, from evolutionary economics, universal Darwinism and even complexity theory, rather than from the meta-narrative’ of Marxist political economy (see, for example, Boschma and Frenken, 2006; Essletzbichler and Rigby, 2007; Martin and Sunley, 2007). It is within this embryonic subfield of evolutionary economic geography that the concept of path dependence has been accorded particular theoretical and empirical significance. Of especial interest is the question of whether and to what extent the evolution of the economic landscape is a path-dependent process, whether the mechanisms that make for path dependence have a quintessentially local dimension in their form and operation, and thus whether, in this sense, path dependence can be seen as a process or effect that is locally contingent and locally emergent, and hence to a large extent ‘place dependent’ (Martin and Sunley, 2006).

However, as we argued in our previous conspectus of the concept (Martin and Sunley, op cit), despite the increasing use of path dependence terminology and notions by economic geographers, there has been little extensive or detailed discussion of what the ontology called for by Scott might look like. In fact, economic geographers have tended to apply the concept of path dependence as if it is self-evident and wholly unproblematic. As Glasmeier (2000, pp. 269-270) complained, in economic geography the concept of

path dependence is often invoked uncritically as an explanation for a particular industrial [and, we might add, regional] experience. Usually lying behind the notion of path dependence is a series of factors that together add up to a directional bias. Just exactly what provokes path dependency is rarely communicated, however; this often erroneously leads to uni-dimensional invocations of the term.

We would go further and suggest that the very idea of ‘path dependent economic evolution’ as advocated by Scott is itself in need of careful interpretation and conceptualisation if it is to serve as the basis of a meaningful ontology of regional growth and development. The stakes here may be high. Path dependence may help explain why regional growth disparities persist; it may help explain why particular industries and technologies develop in certain locations but not in others; and it may help us to understand why some regional economies are better able to adapt over time than others. Our previous paper was intended as an initial exploration of such issues. In that paper, we explored the sorts of
processes that could give rise to path-dependent economic development in a geographical – and specifically, regional – setting, and sought to argue that many of those processes are themselves place dependent. The aim here is to build on that discussion and to focus much more specifically on how far and in what ways path dependence can serve as an evolutionary concept for studying the economic landscape. The thrust of our argument is that this task is problematic. For one thing, we find that despite its declared emphasis on the importance of history, path dependence theory as formulated by its leading architects - David, Arthur, and others - retains elements of equilibrium thinking, which we contend is in tension with the idea of path dependence as an evolutionary concept. We are sceptical that the evolution of the economic landscape ever subtends to any equilibrium state: to be sure, the economic landscape is characterised by self-organisation and order, but these are not the same thing as equilibrium. We argue, therefore, for a wider view of path dependence that allows for patterns and trajectories of development that do not approach or reach an equilibrium state, that do not require an equilibrium interpretation. This leads on to a discussion of how to characterise such paths in a regional context.

2. The Basic Model of Path Dependence and its Geographical Application

One of the abiding intentions behind Paul David’s development and proselytisation of the concept of path dependence has been to persuade economists - and especially neoclassical economists - to move beyond traditional equilibrium modelling and to ‘take history seriously’. This endeavour resonates strongly with evolutionary economists, some of whom, such as Hall (1994), have elevated the notion of path dependence to the status of a ‘first principle’ of evolutionary economics. David (2001, 2005, 2007) himself argues that path dependence is a property of a ‘wide array of processes that can properly be described as evolutionary’, including economic processes. But in what sense, precisely, does path dependence function as an evolutionary concept? Neither David nor Arthur, the primary exponents of path dependence ideas, has set out a fully specified model of economic evolution based on path dependence, nor discussed how the concept relates to other key ideas in evolutionary economics. Nevertheless, it is possible to infer from their various writings the sort of model of economic evolution implied by their depiction of path dependence.
According to Sydow, Schreyögg and Koch (2005), the basic model of path dependence developed by David and Arthur implies a three-stage model of the historical (and, we might add, geographical) development of a technology, industry, institution or organisation. In the first, ‘pre-formation’ phase, as they term it, considerable scope and variety exist for exploring or developing a new technology, product, industry, or institution (and again, we may add, a location for such an activity). In David’s and Arthur’s work the search process is portrayed as largely undirected, and decisions mainly contingent. There may be several alternative opportunities being explored at this stage. A ‘critical juncture’ then occurs – in David’s accounts this is usually an ‘historical accident’ or ‘random event’ - that results in one particular development opportunity or decision being ‘selected’ or preferred over alternatives, a process that David likens to the idea of ‘bifurcation’ used in complexity theory. This opportunity or development then begins to attract other actors, or acquires market influence, and a critical mass around this activity begins to build up and a development path is formed. Once this critical mass achieves a certain size or momentum, the path gets ‘locked-in’, and a third phase of cumulative and self-reinforcing development along this path ensues.

Sydow, Schreyögg and Koch criticise the basic David path dependent model on several grounds. For example, they argue that the assumption that path creation is a random event or ‘accident of history’ ignores the fact that the emergence of a new technology, a new product, or a new organisation, is often the outcome of purposive behaviour and directed decision-making by economic agents (see also Garud and Karnøe, 2001; Martin and Sunley, 2006). Further, they suggest the model is incomplete in the sense that it says nothing about how paths ‘de-lock’, break up and dissolve. These are obviously important aspects of the path dependence idea, and clearly crucial to its relevance as an evolutionary concept. While it is true that David tends to see the emergence of a path in terms of some ‘random event’ or ‘historical accident’, it is not strictly true that he says nothing about how a path ends. In fact he has a very specific view on this, namely that the break-up and dissolution of a technological, industrial or institutional trajectory is brought about by an ‘external shock’ of some sort, which then destabilises the system and opens up opportunities for a new path to emerge. Thus, the basic path dependence model actually posits four stages of the development of a technological, industrial or institutional trajectory: pre-formation, path creation, path lock-in, and path dissolution (see Figure 1).
There are two main ways in which economic geographers have used a basic path dependence model of this sort: to explain the evolution of a particular industry, technology or institution either in a given location (region, city) or across locations. In the former case, interest has focused on identifying the initial locally ‘contingent’ factors or stimuli responsible for the emergence of the industry, technology or institution (the creation of the path) in the area in question, and on the types of self-reinforcing mechanism and co-evolutionary processes that explain its subsequent path-dependent development. In particular, attention is often directed to the role of local ‘network externalities’ (to use David’s terminology) or ‘increasing returns’ (to use Arthur’s). In some applications, particularly in relation to high-tech regions, the local embedding of such network externalities and increasing returns is viewed as key to innovative and competitive success. In such cases, ‘lock-in’ is seen as a positive process. In other applications, however, typically older industrial regions, ‘lock-in’ is seen as a negative feature, whereby a region or locality becomes over-reliant on, or dominated by, a particular self-reinforcing industrial-technological path that renders the regional or local economy increasingly structurally and
technologically rigid, restricting thereby its capacity to adapt to changing competitive forces (‘the weakness of strong ties’ argument). In this version the basic model is often used to argue that the path-dependent overspecialisation of regions is likely to make them vulnerable to sector-specific (and hence region-specific) shocks that then bring about the decline of the industry in question, and with it much of the regional economy as a whole.

In the second type of application of path dependence in economic geographic work, the focus is less on the path-dependent evolution of a particular industry in a particular region, than on how a given industry evolves spatially across a multi-region or multi-location economic landscape (although of course the industry may end up entirely concentrated in just one, particular location). In a similar way that in the basic path dependence model a new technological or industrial path is assumed to be initiated by random, chance or happenstance events, so in this geographical version of the model the parallel assumption is that initially there is spatial indeterminacy as to where a new technology or industry will begin to emerge. Several possible potential locations may be equally suitable initially. In which of these locations the industry or technology is triggered will be largely an accident of history – a situation described by Boschma and van der Knaap (1997) as one in which the ‘windows of local opportunity’ are ‘open’. Which locations then get ‘selected’ for subsequent path dependence development will depend on small arbitrary differences in the conditions and circumstances in those ‘trigger’ locations, and whether and to what extent cumulative self-reinforcing feedbacks emerge:

the discontinuous nature of major innovations ... implies that the spatial formation of new industries involves spontaneity or indeterminacy because it is unlikely to be determined by or bound to particular places... the actual outcome depends on small arbitrary events, magnified by a positive feedback mechanism, which, in our approach, is achieved by the creative ability of firms that build up a favourable local production milieu around them (Boschma and van der Knaap, 1997, p. 182 and p. 187).

The parallel with Arthur’s path dependence model of industrial location is clear:

[E]arly firms are put down by historical accident in one or two locations; others are attracted by their presence, and others in turn by their presence. The industry ends up clustered in the early-chosen places. But this spatial ordering is not unique; a different set of early events could have steered the locational pattern into a different outcome... We might call this view historical dependence. (Arthur, 1994, p. 50).
Thus the eventual spatial pattern of an industry is interpreted as being the outcome of early (chance) events and subsequent spatially selective path dependent cumulative processes. Models of this sort have tended to focus overwhelmingly on how new industries emerge and develop across space, and have had little to say about how a given spatial-industrial path dissolves, though the implication of the windows of locational opportunity concept is that each new technological innovation opens up new windows and hence new spatial configurations of economic development, so that “the long term evolution of the spatial [economic] system is potentially unstable” (Boschma and van der Knaap, 1997, p. 198), How this relates to the observation that long-run patterns of relative regional prosperity are often highly persistent (and path dependent) over long periods of time is not clear.

Whilst these uses of the basic path dependence model in economic-geographic work are certainly suggestive, they leave several issues still largely unresolved (see, for example, Table 1). Many of these have to do with how we conceptualise the meaning and nature of path dependence within regional and local settings. Others relate to the sources of path dependence and how far these are shaped by local conditions and circumstances, that is how far path dependence is itself place dependent. Still others concern why the degree of path dependence seems to vary across the economic landscape. And, importantly, there are basic questions about what sort of economic evolution is implied by path dependence: does it imply slow, incremental change and development, or a more punctuated pattern, of successive periods of relative stability of economic structures and development forged by episodic bouts of radical industrial and technological change? It is on this latter issue that we wish to focus in what follows.3 This turns out to be a far from straightforward task.

3 In focusing on the role and meaning of path dependence as an evolutionary concept, we by no means wish to imply that the other issues concerning the notion of path dependence set out in Table 1 are secondary or unimportant: far from it. But arguably the role of path dependence as an evolutionary concept is the most fundamental. For a discussion of the other issues raised in Table 1, see Martin and Sunley (2006).
Table 1: Some Key Issues in the Application of the Concept of Path Dependence in Economic Geography

<table>
<thead>
<tr>
<th>Key Questions</th>
<th>Issues</th>
</tr>
</thead>
<tbody>
<tr>
<td>What is the object of study?</td>
<td>What precisely are we referring to: a region’s firms, its industries, or the regional economy as an aggregate? What is the relationship between the micro-level and the regional aggregate level? Path dependence is multi-scalar process, operating at different levels. How do these interact within regional economic spaces? Is regional path dependence an emergent effect?</td>
</tr>
<tr>
<td>What are the mechanisms of regional path dependence?</td>
<td>What are the sources of path dependence within local and regional economies (sunk costs and infrastructures, external economies of industrial localisation, general agglomeration economies, local socio-cultural-institutional embeddedness, local and extra-local economic dependencies and linkages)? How do these mechanisms vary (and interact) across space?</td>
</tr>
<tr>
<td>Do all parts of a regional economy display path dependence?</td>
<td>Is a certain threshold of interaction and inter-relatedness within a region economy required before it displays aggregate path dependence? Can different paths co-exist within a region, and how do these interact – what is the nature of ‘inter-path coupling’?</td>
</tr>
<tr>
<td>What are the sources of regional path dependence?</td>
<td>Some will be more or less specific to particular types of regional economy; but, different forms of path dependence are likely to co-exist within a regional economy. In this sense, the notion of regional path dependence is a complex, multi-dimensional. Is a single, overarching theory possible?</td>
</tr>
<tr>
<td>Do the nature and strength of path dependence vary from region to region?</td>
<td>Regions differ markedly in economic structure, institutions, connections to other regions and beyond, etc, so the nature and degree of path dependence might be expected to vary from region to region. In what ways is path dependence a locally contingent - that is, a place dependent - process?</td>
</tr>
<tr>
<td>In what sense does path dependence explain the evolution of the economic landscape?</td>
<td>What sort of evolution is implied by path dependence? Is path dependence consistent with incremental economic change and mutation? Or does it imply a ‘punctuated equilibrium’ process of evolution and transformation of the economic landscape? Where do new paths come from, and why do they emerge where they do?</td>
</tr>
<tr>
<td>How pervasive is path dependence in the economic landscape?</td>
<td>Is it a fundamental feature of regional development and evolution – ie inevitable and indeed necessary for regional growth to take place? Or is it more typical of economically specialised regions and localities, and less likely to emerge in areas with diverse economic-technological structures?</td>
</tr>
</tbody>
</table>

Source: Based on Martin and Sunley (2006)

quintessential feature of this ‘equilibrist methodology’ (as Setterfield, 1997, calls it) is that the development of the economy is interpreted not as being shaped in any significant and persisting way by particular events that occurred in the historical past, but as a movement towards a hypothetical equilibrium outcome.
The aim is to demonstrate how, under specific given (and typically highly abstract, simplified and idealised) assumptions (as to consumer tastes, technological knowledge, the nature of competition, institutional arrangements, etc), an economy tends ineluctably and deterministically towards a limiting unique \textit{ex ante} equilibrium state that is invariant over time and space. The economy, in other words, is conceptualised as an \textit{equilibrium process}, which (following Harris, 2004) might be depicted as:

\textbf{Equilibrium Process:} $x(t+1) = F^e_x(x(t)), \ -\infty \leq t \leq +\infty; \ x^e = \text{equilibrium point}$

where $x$ is the economic state or outcome of interest, and the function $F^e_x$, governing the change over time of the system, generates a unique and stable equilibrium $x^e$. In such a conception, equilibrium is not a real emergent outcome of actual historical processes and events, but an abstract solution state determined by the specific assumptions, equations, and exogenous parameters and variables built into a formal economic model, the basic purpose of which is to determine the existence and stability of equilibrium. Any notion of ‘dynamics’ in such models concerns only the model economy’s movement towards its equilibrium: once in that equilibrium state, the system is in stasis. Equilibrist economics is basically antithetical to notions of history and evolution (Kaldor, 1934, 1972; 1985; Robinson, 1974; Setterfield, 1995, 1997; Harris, 2004).

In contrast to equilibrium economic processes, in an economy characterised by path dependence the specific details and sequence of historical events govern the unfolding course of development – what David calls ‘historically contingent evolution’. Following Harris (2004) and Page (2006), such a process might be expressed as:

\textbf{Historical Process:} $x(t+1) = F_x(t)(h(t)x), \ -\infty \leq t \leq +\infty,$

where $h(t)x$ is the history of past outcomes of $x$ from $t=0$ up to time $t$, and the function $F_x(t)$ maps that history into the next outcome. The outcome function

\footnote{While there are many different definitions of such an equilibrium, according to Setterfield (1997), an equilibrist methodology of this sort is typically characterised by two distinctive features: first, the specification of a model of structural equations conditioning endogenous variables on exogenously given ‘data’ – usually a set of variables and coefficients whose values are imposed on the system from without; and, second, the construction of such models so as to yield stable equilibria, that is points to which the system will return following some initial displacement.}

\footnote{The historical sequence, $h(t)x$, could also include the past values of other external factors, say $y$, that have also shaped the development path of $x$.}
can itself change over time so it is indexed by \( t \). In such a system, the present state of the economy will depend on where it has come from, and on how it got there: this is what is generally meant by path dependence. There is nothing inherent in such an historical process that necessitates that it possesses or reaches a stable equilibrium state. Indeed, the concept of path dependence can be argued to be fundamentally antagonistic to an equilibrist methodology. According to the latter, the long-run equilibrium state can be defined and reached independently of the path taken towards it, whereas with path dependence any long-run configuration that is reached by the economy will depend on the path take towards it. There is thus no predetermined economic outcome, no outcome independent of history or context. To argue that an economy, an economic landscape, is a path-dependent historical process would thus seem to be incompatible with simultaneously arguing that it is an equilibrium process.

Yet, whilst emphasising the need to move beyond the equilibrist methodology of mainstream economics, Paul David and other path dependence theorists nevertheless seem reluctant to relinquish the idea of ‘equilibrium thinking’ altogether. In fact, recently, David has explicitly referred to his approach to historical economics as ‘path dependent equilibrium analysis’ (David, 2005, p. 153). His strategy to reconcile the apparent contradiction in this phrase is to define path dependence in terms of the dynamics associated with particular types of non-ergodic stochastic processes and systems that possess a multiplicity of limiting distributions, that is multiple equilibria:

The elaboration of theories around the core concept of path dependent dynamics... encourages and enables economists to entertain the possibility that, in place of a unique equilibrium-seeking dynamic, they should envisage a process that is seeking an

---

6 For example, the same seemingly irresistible attachment to equilibrium is to be found in Page’s (2006) otherwise very useful attempt to distinguish between different forms and degrees of path dependence. He states: “of course a process need not attain an equilibrium distribution or outcome, but for the purposes of this essay I restrict attention to that case, with one exception” (p. 92).

7 More explicitly, David defines path dependence in terms of non-ergodic Markov chain models that possess two or more absorbing states. An absorbing state is one from which there is a zero probability of exit or movement to another state. In such models, the limiting distribution depends on the initial distribution, that is on where the system started: different starting distributions will result in different limiting (equilibrium) distributions: hence the idea of multiple equilibria. We should note here, however, anticipating the discussion that follows, that absorbing Markov chains are only one form of non-ergodic process, and that in general the latter do not have to possess any long-term limiting distribution or stable equilibrium state.
evolving and historically-contingent equilibrium (David 2005a, p. 2).

Under such conditions,

Small events of a random character – especially those occurring early on the path – are likely to figure significantly in ‘selecting’ one or other among the set of stable equilibria, or ‘attractors’ (David, 2007, p. 151).

Which of these multiple equilibria is reached or ‘selected’, it is contended, will depend on the initial state of the system – on initial ‘random events’ – and the chains of transitions produced by repeated iteration of the system over time. In this sense, there is no single *ex ante* unique distribution of the economic system, as in general equilibrium economics: the actual limiting distribution depends on history, and particularly where the economic system (for example, the technological or industrial structure) started. Further, according to David, once the processes of path dependence have locked the system (technology or industry) into one of these alternative stable equilibria (or ‘attractors’), it requires an external shock to break the particular development path or trajectory in question. The evolution of the technological, industrial and institutional structure of an economy in this path dependence model would thus seem to be akin to ‘punctuated equilibrium’, a (stochastic) pattern of historically contingent, path dependent evolution in which particular paths of technological, industrial and institutional development become ‘selected’, locked-in to stable equilibria, and, at some point, eventually disrupted and broken up by external shocks. As David puts it:

Sudden shifts in structure, corresponding to the new evolutionary biologists’ notion of ‘punctuated equilibria’... may open up a way for the formulation of dynamic models that are compatible with ‘stage theories’ of development (2005, p. 187).

What is interesting about this ‘path dependent equilibrium analysis’ is that it resonates with how notions of ‘path dependence’ and history have been used in the so-called ‘new economic geography’ (NEG) that has arisen over the past two decades. Adherents of this approach to analysing the space economy are explicitly equilibrist in orientation, asserting that it is “the general equilibrium modelling of an entire space economy which sets the approach apart from traditional location theory and economic geography” (Fujita and Mori, 2005). Moreover, whilst committed to this equilibrist representation of the economic
landscape NEG theorists simultaneously claim to recognise that ‘history matters’ and that their models incorporate ‘path dependence’. The way this is achieved is precisely by constructing (deterministic) models that yield multiple equilibria outcomes (in this case equilibrium spatial patterns of industry and employment) depending on the ‘initial conditions’ (such as the level of transport costs, the relative mobility of labour and capital, the initial distribution of industry between regions, and the number of regions) specified in the models. For example, the basic two-region ‘core-periphery model, the foundation of NEG theory, possesses five such possible equilibria (three stable and two unstable). Because the ‘initial conditions’ determine which of the alternative equilibrium landscapes emerges, NEG theorists feel able to claim that ‘history matters’ in their models. And because these models generate their equilibrium outcomes via processes of self-reinforcing agglomeration processes, it is also claimed that they embody path dependence (or ‘locational hysteresis’).

Now the notion that ‘initial conditions’ matter is certainly an improvement over the timelessness of conventional equilibrium economics, since potentially it at least draws attention to one aspect of the past history of a system in the determination of future outcomes. However, to our mind, characterising path dependence as a problem of multiple equilibria (whether stochastic, or deterministic as in NEG models) is to restrict the concept and its relevance for analysing the evolution of regional and local economies. In David’s work, as in NEG models, multiple equilibria are conceived as representing a system that possesses a variety of ‘locally stable attractors’ (to use his term), from which the final equilibrium position is ‘selected’ on the basis of the system’s initial starting position or state. In effect, if we know the various alternative possible starting states, the array of possible final outcomes to which they may conform becomes determined a priori, and all that remains to be ascertained is which outcome of the possibles will actually be selected. Thus

Characterising path dependence as a problem of multiple equilibria constitutes a limited conception of this phenomenon because it overlooks the possibility of a path dependent process creating its own set of final outcomes in the course of its evolution. The basic distinction here is one of ontology. In a selection process, the environment comprises a pre-prescribed or given external reality to

---

8 The literature on NEG models is now extensive, but for representative expositions see Fujita, Krugman and Venables (1999), Baldwin et al (2003), Henderson, (2005) or Brakman, Garretsen and Mearrewijk (2008). For a recent assessment of these models from a proper economic geography perspective, see Martin (2009).
which decision makers must adapt (through a process of selection). In a creation process, however, the environment is not pre-prescribed, but is instead ‘open’ – it remains to be constructed and defined in the course of the concrete functioning of the system” (Setterfield 1997, p. 64, emphasis as in original).9

What is being argued here is that even the idea of a path-dependent equilibrium would seem to run counter to the basic principles associated with historical time. However an equilibrium is defined, it is a state from which a system will display no endogenously generated tendency to deviate. External shocks, as in David’s schema, may disrupt and destroy that equilibrium, but these are extraneous influences. What this suggests, then, is that once we are in equilibrium, history and change effectively end: the future is predetermined by state of the system corresponding to the equilibrium that has been achieved. To define path dependence as the historically contingent selection process between multiple equilibria seems to imply that path dependence only matters in the selection and movement towards an equilibrium position, and that once in that equilibrium position path dependence merely serves to reproduce that equilibrium, that is to entrench stasis, a position of no change.10

Setterfield’s discussion of the relationship between path dependence and equilibrium is one of the most detailed we have been able to find. However, although he argues forcefully for the incompatibility of history and equilibrium, even Setterfield cannot bring himself to dispense with equilibrium thinking. He suggests – unsuccessfully in our view – that the notion of equilibrium can still be usefully retained in an historical or path dependence approach to economic theorising under two possible strategies (Setterfield, 1997). The first is to treat equilibrium in the economy as a ‘pedagogic device’ to ‘lock-up without ignoring’ the complexities of historical time. But surely this is to introduce a form of conditional closure to what in is reality an open, historical system, and seems to

---

9 Although Setterfield makes this argument in the context of deterministic systems, it applies equally to stochastic systems like the absorbing Markov chain processes that David uses to conceptualise path dependence.

10 Of course, it could be argued that equilibrium economics has long moved on beyond traditional equilibrist methodology, and while retaining its focus on the existence and achievement of equilibrium, has come to embrace the idea that history matters through the construction of models of sequence economics, temporary equilibrium, and non-latonnment processes. But as Setterfield retorts, these extensions actually constitute “nothing more than glorified conventional disequilibrium analysis and are roughly in keeping with equilibrist methodology... the language of the analysis me be that of sequences and temporary equilibria, but the substance does not involve path dependency, as the economy is held ultimately to converge to a predefined, ‘fully adjusted’ equilibrium position” (op cit, p. 62)
us to be little different from the position adopted by mainstream equilibrium economics. In effect it is to assume that the various mechanisms and forces that produce change in the economy are ‘held constant’. As Setterfield himself states:

The pedagogic use of equilibrium, designed to ‘lock up without ignoring’ pertinent features of historical time, suggests that in some circumstances it may be strategically useful to demonstrate what it would mean for the economy to be in ‘equilibrium, even if historical motion and change are believed to be pervasive (op cit, p. 68).

This is difficult to reconcile with an interest in economic evolution if indeed that is the focus of our interest. His second suggestion is to treat equilibrium as the actual ‘temporary’ outcome of a path dependent process that may yet give rise to a subsequent endogenous process of ‘innovating out’ of equilibrium. Path dependence then is defined in terms of a sequence of ‘temporary equilibria. Setterfield contends that if all equilibria are conceived as essentially temporary equilibria, the “the antagonism between historical time and the concept of a state of equilibrium is at least attenuated” (op cit p. 68). While this idea of ‘endogenously innovating out of equilibrium’ is a useful corrective to the usual appeal to the necessity of an exogenous shock to move the economic system out of an equilibrium state, we still fail to see why the notion of a ‘temporary’ equilibrium is needed. How do we (or economic agents themselves) know an economic system is in a state of ‘temporary’ equilibrium? And how short in duration can a position of ‘temporary’ equilibrium be before it becomes indistinguishable from a process of incremental endogenous change?

Indeed, the very idea of equilibrium – whether path dependent or not – is difficult to reconcile with what is generally taken to be the defining feature of an economy viewed as an evolving system. According to Witt (2003, 2006) for example, the key focus of evolutionary economics is on the processes and mechanisms by which the economy self-transforms itself from within, that is on processes of endogenously generated change. This conception is fundamental to what Castellacci (2006) calls the evolutionary ontology:

The co-existence of random and systematic factors driving economic evolution ... and the combination of inertial and dynamic forces, both constitute important elements in an attempt to explain the most important stylised facts about economic evolution.... the

---

11 As Schumpeter (1944) insisted, transformation arises from within the socio-economic system, and adaptive development is the primary process by which this occurs.
existence of structural change (the old Schumpeterian ‘creative
destruction’), persistent differences in growth rates between
regions and countries, phenomena of path dependency, and
cumulative causation patterns. Such real phenomena are regarded
as unique events in historical time. Differently from the
neoclassical metaphor of a steady state, evolutionary scholars
describe an ever-changing and never-ending process of growth and
transformation (p. 869).

Evolutionary thinking requires a shift in mindset from the characterisation of
equilibria to the specification of dynamic and historical processes. Economic
evolution is a species of ‘far-from-equilibrium’ process and what keeps
economies far from equilibrium is the particular set of knowledge-generating
and application processes that define a modern economy. Since knowledge is
never in equilibrium, so an economy can never be in equilibrium (Ramlogan and
Metcalfe, 2006), and its behaviour cannot be understood by asking whether it
has any equilibrium or steady states (Durlauf, 1997).12 Both David and Arthur
(and other scholars of path dependence) recognise the role of learning as a causal
mechanism of path dependence, but at the same time operate with an
equilibrium interpretation of path dependence. Further, complexity theory tells
us that ‘far-from equilibrium’ systems can be highly structured and patterned,
and strongly path dependent.13 In such complex systems, structures and patterns
emerge from processes of self-organisation not as outcomes of asymptotically
equilibrating mechanisms.14

It would seem, then, that we can distinguish at least three conceptions of
path dependence as an evolutionary concept. The first and most restrictive is the
David-type conception that defines path dependence in terms of the historically
contingent selection of, and lock-in to, one of a multiplicity of possible stable
equilibrium outcomes. Once locked in, it requires an external shock of some kind
to dislodge or de-lock the equilibrium state. The process of economic evolution
implied is akin to ‘punctuated equilibrium’, wherein a technology or industry – or
regional economy – becomes locked in to a particular selected path or state, in
which it remains unchanged (in ‘equilibrium’) until such time that an external

12 Ramlogan and Metcalfe make the point that if knowledge ever did reach an equilibrium state
then all economic change and development would cease.
13 Complexity theorists usually use the term hysteresis rather than path dependence. Though not
identical, the two notions are related.
14 The rise of agent-based models within some quarters of neoclassical economics can be
argued to be a manifestation of recognizing the value of assigning more importance to economic
process than to end states (see, for example, Axtell, 2007).
shock disrupts that state (Figure 2). It is perhaps significant that all of the primary examples of path dependence cited by David (the QWERTY keyboard, VCR video, AC electrical current, light-water reactors) have been of technologies or technological configurations that once 'locked in' remained largely unchanged. These might be interpreted as examples of 'stable equilibrium' states, but they are clearly only one type of economic evolution, and arguably a restricted form at that.

The second conception is of the Setterfield type, wherein path dependence processes generate a ‘temporary equilibrium’ outcome that then gives rise to a subsequent endogenous process of ‘innovating out’ of equilibrium. According to Setterfield, the very fact of the economic system in question (an industry, say) being in a state of ‘temporary equilibrium’ itself will tend to stimulate purposive behaviour by some economic actors to explore pre-emptive breaks from the locked-in technological-industrial activity in order to establish a new competitive ‘temporary equilibrium’. Path dependent economic evolution in this schema is one of a succession of ‘temporary’ equilibria’ (Figure 3).
A third conception is that which frees the idea of path dependence from any necessary connection with equilibrium, and which views path dependence as a dynamic open historical process by which technologies, industries and institutions evolve along unfolding trajectories: what we might term a non-equilibrium conception of path dependence (Figure 4).  These trajectories are shaped - to some extent at least - by the sequences of prior developments and influence of earlier events, but also by the evolution of the processes (mechanisms) of path dependence themselves. This is what the time-varying historical process function $F_X(t)$ used above is meant to capture. Under this approach, explicit allowance is made for the possible interaction between the evolving technology, industry or institution and the various path dependence processes (such as learning, network

---

15 We prefer the phrase 'non-equilibrium' here, rather than 'disequilibrium' or 'out-of-equilibrium', since both of the latter continue to imply that there is some latent equilibrium outcome to which path dependence would lead if only the system in question were free to do so. The term 'non-equilibrium' is intended to imply that no such latent equilibrium may exist at all.
Figure 4: An Open, Non-Equilibrium –Type Model of Path Dependent Evolution

Time

possible eventual maturation, exhaustion and decline of path

externalities, etc) that are shaping it. Put another way, a path-dependent technological, industrial or institutional trajectory and its associated path dependence processes may co-evolve. Such a system may never approach any form of equilibrium. Furthermore, such a schema allows for various possible evolutionary pathways. It can encompass the case of endogenously generated incremental path dependent evolution as well as externally-initiated punctuated forms. It allows for the adaptation and mutation of a technology, industry or institution over time (of which numerous examples abound – indeed most technologies, industries and products are characterised by this mode of evolution). And it also includes those cases where a path-dependent form of development endogenously generates the mechanisms of its own relative decline and dissolution.

These different possibilities seem to us to be highly relevant to discussions of the path dependent evolution of local and regional economies. Thus, for example, in some regional and local contexts we might observe instances where the local path dependent development of an industry follows the upper evolutionary trajectory in Figure 5, and the path dependent lock-in of the local economy to a particular industrial specialism eventually becomes a source of inflexibility, vulnerability and decline. This decline can itself be a slow process (for example due to a progressive loss of relative competitiveness because of insufficient or only sluggish improvements in labour productivity), and need not
necessarily be a sudden collapse due to an external shock (as in David’s rendition of the end of an industrial/technological path). In other instances, we might observe the sort of trajectory depicted in the lower part of Figure 5, whereby a local industry succeeds in ‘reinventing’ and reconfiguring itself over time, thus avoiding the problems of lock-in to an increasingly rigid or uncompetitive structure. In many regions, of course, both sorts of sectoral path dependent evolutionary trajectory could co-exist, and may even interact, in positive or negative ways. Releasing the notion of path dependence from any necessary connection with equilibrium states, then, raises a wide set of issues about exactly how we understand evolution along an industrial or technological path.

![Figure 5: Alternative Models of the Path Dependent Evolution of an Industry or Technology](image)

4. A Non-Equilibrium Approach: Evolution along Paths

As we have seen, the notion of path dependent equilibrium is based on the view that the operation of feedback processes gradually increases the probability of a certain (equilibrium) outcome. These processes cumulatively reinforce the
direction of change. In contrast, however, an open non-equilibrium interpretation of path dependence needs to recognize that while institutions and past trajectories channel change, at the very same time they also enable adaptation and the launch of new paths. Indeed, in some cases the very same mechanisms summoned by equilibrium path dependence – increasing returns, learning, network externalities and the adoption and spread of technological standards – are precisely those that are acting to maintain and even increase the probability of adaptation and departures to new paths. We would contend then that a realistic theory of path dependence needs to examine the conjunction of both sets of constraining and enabling effects.

This non-equilibrium perspective has important implications for our understanding industry and product life-cycle approaches that attempt to explain evolution along particular paths. According to product cycle theory, as products become more standardized and routinised, their production location shifts from creative seed-beds and core cities, where product innovation first occurs, towards lower labour cost locations in more peripheral regions and countries (see, for example, Vernon, 1966). The predictable evolution along a path determined by the maturity of the product thus leads to the dispersal of production and growing wealth in lower cost regions. This teleological evolution towards types of spatial equilibria is, of course, highly problematic (Storper, 1985; Markusen, 1985; Taylor, 1986). In Storper’s (1985) evaluation, for example, the product cycle approach was too essentialist as it sought regular mechanical, and closed system relationships among variables and failed to recognize the full open-ended ontology of economic processes. In particular, he suggested that industrial ‘life cycle’ notions extrapolate from transitory and contingent empirical phenomena and credit them with the status of a teleological developmental logic. The choices in manufacturing and business strategy envisaged are too determinate and too closely related to the extent and maturity of the market: “the cycle conceives of technological history as a series of determinate, repeating events, which it assumes to be a common to all industries. Thus it would turn economic history into a kind of ‘natural history’” (Storper, 1985, page 271). Storper advocated a more open-ended view of history as a series of unfolding events and in this way he dismissed evolution as an endogenous organic unfolding and instead gestured towards a more neo-Darwinian view of industrial change. However, while his argument leaned towards a more idiographic and unpredictable understanding of industrial and
regional paths, a completely open-ended approach risks losing sight of path dependence altogether.\textsuperscript{16} More recent explanations of industrial life-cycles certainly identify generic forms of path dependence in a range of industries. There are several different varieties of the industry life-cycle approach but they share a similar schematic interpretation of how industries develop (Agarwal and Gort, 1996; Agarwal, 1998; Klepper, 1997). Most accounts distinguish between an early formation stage, an intermediate development stage and a mature stage. During the formative stage firms enter a new industry by producing a relatively new product and face a great deal of uncertainty. During the second stage process techniques become more refined and markets and outputs grow rapidly. During the mature stage, production techniques are further refined, market growth levels off and significant innovations are fewer. In terms of population dynamics, the entry of new firms is greatest during the formation stage. Typically it is argued that a sharp drop, or shakeout, in the number of producers occurs during the development phase and the number of firms continues to decline in the mature stage (Klepper and Simons, 2005). Klepper (1992) argues that innovation also tends to evolve during the course of an industry’s life-cycle. Innovative activity is greatest during the early formation stage and new and small entrants have an innovative advantage. In the development phase, there is a reduction in innovative activity and a shift towards established large enterprises. In the mature phase, any new entrants face a comparative technological disadvantage. Audretsch and Feldman (1996) spell out some geographical implications and argue that the propensity for innovative activity to be geographically clustered is strongly related to the stage of the industry life-cycle. Innovative activity tends to be more clustered in the early and formation stages as small new entrants benefit greatly form localized knowledge spillovers.

The stylized facts of industry life cycles suggest that path dependence arises from several different forces (Klepper 1996; 2002; Klepper and Simons, 2005). In one approach a radical invention induces firms to enter a new industry. Those that are successful innovators expand their production and the pressure of output growth on prices causes the less successful innovators to exit the industry. In another approach, firms enter an industry with alternative and competing product designs. During the course of time one design becomes the ‘dominant’ or

\textsuperscript{16} It is analogous in a sense to Paul Davidson’s (1991) post-Keynesian economics in which decision-making is so non-ergodic that the past is no real guide to the future.
standard design. Subsequently, opportunities to enter the market based on new designs are diminished and firms compete on the basis of process innovations. In other words, the lock-in of the dominant design reduces opportunities to enter (Klepper, 1996). In the third approach, described by Klepper (2002), prices tend to fall as output in an industry expands, and this limits the ability of later entrants to catch-up with the size of earlier entrants. But larger firms benefit most from R and D as they can apply its benefits to the largest amounts of output. As price continue to fall, therefore, the smallest firms and least able innovators are driven out of the market. While this debate continues, it is clear that all these varieties of explanation rely on forms of path dependence. Whether it is the lock-in of a dominant design through network externalities and learning effects, or the positive feedback from R and D and learning-by-doing advantages to first-movers, these dynamics envisage types of cumulative process. They all suggest that the sequencing of firm entry and the accumulation of capabilities through time are crucial to industrial evolution, and hence to the path dependent development of local and regional economies.

A non-equilibrium interpretation emphasises that such cycles are not deterministic. In the first place, industrial trajectories are not exclusively determined by their product market maturity. For example, Hudson (2005) argues that the trajectories of industries in North East England were ‘path contingent’ as they depended not only on trends in international markets but also national political decisions. Second, learning by experience and the accumulation of resources actually facilitate actors’ reflexive adaptation (Rantisi, 2004). As a consequence, important departures from the life-cycle model have been found in some industries as new markets and niches are repeatedly constructed and exploited. For example, Klepper and Thompson (2006) suggest that the continual rise in the number of entrants in the laser industry may be due to the importance of specialized submarkets. The proliferation of submarkets provides new niches for entrants and, as the nature of the technology does not allow significant economies of scope, established incumbent firms can not dominate these new markets. Where new entrants can occupy distinctive market niches and differentiate themselves from incumbent firms they prove resilient and sustainable. Further evidence of enabling forms of path dependence is that product innovation does not always wane in the mature phase of certain industries. Instead, it may rise sharply. For example, in the mature automobile industry, Japanese and some European producers gained an innovative lead over American producers through both process and product innovations (Klepper,
Such cases of ‘resurgent innovation’ have been widely described as evidence of a phase of ‘dematurity’ (Ibid: Storper, 1985). Such developmental patterns would seem to fit the path dependent model in the lower part of Figure 2, whereby an industry (and that part of a local or regional economy dependent on it) undergoes radical adaptation, in effect renewing (or extending) the industry’s development path.

In order to explain why some mature sectors succumb to negative lock-in and complacency while others demonstrate increased innovativeness, it is crucial that we examine the intersection of enabling and constraining processes. For instance, cognitive path dependence means that cognitive priorities and frames are typically less than fully rational and tend to be inductively based on past experience (Beinhocker, 2006). There is some evidence that most firms pursue path dependent learning (Chapman, 2005; Glasmeier, 2007), and in many cases more radical innovations are pursued and developed by smaller start-up firms. However, in some cases, new firms may also stimulate incumbent firms to diversify and adopt new technologies (Malerba et al, 1999). Given this, the stylized general paths of industry evolution show the outcomes of a continuous tension between channeling processes and those that yield concrete variations and departures from standard life-cycle patterns. Furthermore, another source of non-equilibrium tendencies may be the place-specific interactions between different industrial paths.

In the case of a region or locality dominated by a single industry, then the product and technological maturity of the industry will obviously play a key role in shaping the path of the relevant geographical economy. However, in all probability such a simple and direct influence is likely to be unusual, for several reasons. First, as several authors have noted, even in the case of a single industry prone to clustering it has been found that different clusters show different paths and life-cycles (Menzel and Fornahl, 2007). For example, Saxenian (1996) showed that, due to variations in local firm strategies and network structures, the Boston Route 128 computing cluster declined just as the Silicon Valley cluster expanded. Second, the translation of industrial maturity into cluster and regional development is likely to be complicated by the heterogeneity of responses of different firms, and by cross-scale interactions between firms, clusters and urban and regional contexts. Where firm adaptation is short-termist and primarily aimed at cost-cutting it may effectively undermine the adaptability of larger scale entities
In such cases, there are local conditions and place dependencies that shape the competitiveness of different clusters so that an understanding the state of the industrial life-cycle is necessary but not sufficient. Moreover, most high-technology clusters are actually composed of several technological trajectories and the switching of resources between them may be especially important in times of crisis (Bathe et al., 2001). Most life-cycle approaches, however, do not consider the possibility that the co-location of different industries in a particular region makes a real difference to the co-evolution of their trajectories. It is entirely possible, of course that the different industries and activities that make up a given region’s economy evolve wholly independently of one another. In principle, this could well occur, and in practice it is likely that certain of a region’s industries evolve in a path dependent manner but are uninfluenced by the development paths of other local industries: a situation which we have called ‘multiple unrelated path dependence’ (Martin and Sunley, 2006).

In reality, however, economies are typically ensembles of sectors in which productivity growth is linked by income and expenditure flows (Metcalfe et al., 2006). In a local or regional context there may well be interactions between industrial paths through, for example, upstream and downstream input-output linkages, knowledge spillovers, labour pooling, positive service and infrastructure externalities, or, conversely competition for land, finance and skilled labour. In other words, it is more likely that there is at least some degree of ‘multiple related path dependence’ across an urban or regional economy (Martin and Sunley, op cit). In a different context, Bassani and Dosi (2005) argue that systems may be composed of path-dependent entities but whether or not the system as a whole will be path dependent depends on the structure of the interactions between the constituent entities. They contend that when such interactions are strong - above a certain ‘threshold’ - path dependence among the constituent entities induces path dependence.

\[\text{\textsuperscript{17}}\text{In evolutionary economics, adaptation refers to a process of response to the selection environment so that adapted entities fit the environment and are apt or fit for purpose (Metcalfe, 2005). Adaptability on the other hand, refers to “the potential to adjust to changing circumstances in an appropriate way: it is about the capacity to respond to changes in the selection environment: to maintain good design” (Ibid, page 414). However, economic adaptation remains a layered and contested concept: does it refer, for instance, to the process in which entities ensure that they are fit for market survival or to the process of using resources appropriately and satisfying needs, as clearly the two may overlap but are not synonymous? It is important therefore to clarify the precise firm strategy taken in response to a specified environment and to examine its implications for adaptability.}\]
dependence at the level of the ‘aggregate’ system, although as they acknowledge, there is no a priori way of specifying what that ‘interaction threshold’ has to be before path dependent outcomes become observable. Grabher (1993), of course, argued that in the case of the Ruhr’s iron and steel complex strong interlinkages had produced system-level (region-wide) path dependent behaviour. In most regions there will be groups of inter-related or complementary industries and activities, linked either by direct input-output relationships, or by various indirect (or untraded) interdependencies and externalities. That is to say, various networks and structures of inter-relatedness can emerge between different sectors and activities within a region, thus suggesting the possibility of what we might call ‘path-interdependence’, that is situations where the path dependent trajectories of particular local industries are to some degree interdependent. The extent and significance of this inter-linking path effect is a key issue for further research.

Evolutionary economists increasingly emphasise how path dependence involves the co-evolution of different ‘arenas’ – such as the economic, technological, institutional and socio-cultural. And some geographers have shown how local path dependence in specialised clusters also occurs through such mutually reinforcing co-evolution of local economic, technological and socio-institutional ‘arenas’. But path dependent development may also involve the complex co-evolution of overlapping sets of inter-related industries, and thus may be just as important in shaping the evolution of economically diverse regions as it is in highly specialised ones. At present, however, there is a dearth of research into these regional co-evolutions so that we do not know how far they can effectively change the course of an industrial path, possibly by stimulating adaptation and rebalancing of the local industrial structure, once again disrupting any tendency to equilibrium.

5. Beyond Equilibrium: Place and Path Creation

Partly as a consequence of its focus on evolution along particular paths, and its preoccupation with movements to equilibrium, the path dependence literature has said little about the radically disequilibrating phenomena that are the creation of new technological and industrial paths. As we have seen, David’s equilibrium-type model sees path dependence as primarily constraining and therefore has no option but to ascribe the appearance of new paths to happenstance events or exogenous shocks to the system. But if path dependence can, in some circumstances, also be enabling then this gives us a much wider range of possible cases of path creation (see Figure 6). Path
creation varies both in terms of whether past trajectories are enabling or constraining and in terms of whether origins are intended and deliberate or accidental. As Garud and Karnøe (2001) argue, we should not underestimate the role of purposeful, deliberate and strategic human action (also Schienstock, 2007). In their view, entrepreneurs are embedded in paths but not completely constrained by them; indeed they demonstrate ‘mindful deviation’ from prevailing ideas and accepted ways of doing things. However, it is important to realise that path creation will inevitably involve a complex admixture of deliberate agency and accidental and unintended emergence. Meyer and Schubert (2007), for example, argue that the formations of all technological paths lie somewhere on a continuum between completely accidental and random emergence and deliberate intended creation. Agents are aware of emerging paths and invest in and bet on them, but the results of their actions are not always those intended. The constitution of new paths may be ‘accidental’ then in the sense that deliberate actions often have much longer and wider consequences than the actors intended (David, 1999a, Puffert, 2001). To return to Figure 6, what we intend to show is that while the classic model of path dependence is rooted in box 4, more recent work in economic geography has adopted positions in boxes 2 and 1 and has put much more emphasis on the re-use and transfer of resources and competences.
The Place of Path Dependence

### Figure 6: Varieties of Path Creation

<table>
<thead>
<tr>
<th>Path and place effects</th>
<th>Origins of new path</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Deliberate and intentional</td>
</tr>
<tr>
<td><strong>Enabling new paths</strong></td>
<td>1. Agents search for opportunities, re-use resources, transfer competences as basis of new growth</td>
</tr>
<tr>
<td><strong>Constraining to existing path</strong></td>
<td>3. Designed interventions to break path or switch location to overcome lock-in</td>
</tr>
</tbody>
</table>

This wider set of possibilities raises a series of questions about the strength of path dependent enabling effects. The dominant view in evolutionary economic geography has been that pre-existing paths of development are all but irrelevant in determining where new industrial paths emerge and become established. For example, Storper (1999) proposes the so-called ‘window of locational opportunity’ approach’ outlined above, in which new technologies start as generic assets and then only subsequently evolve into specific assets. Initially, new sectors based on radically new technologies have few established specific inputs so that they ‘invent’ their own input chains and the associated knowledge, which is why, he claims, their initial location is basically serendipitous: this explains why, he argues, the semiconductor industry grew up in Silicon Valley and not alongside its parent industry – radio and television equipment – on the East Coast. New industries enjoy moments
of enhanced locational freedom called ‘windows of locational opportunity’ (Scott and Storper, 1987). In such periods, “Capitalism is capable of escaping the past to create new localizations of industry” (Storper and Walker, 1997, page 71). Similarly, according to Boschma and Frenken (2004, pp. 20-21),

the evolutionary approach argues that the selection pressure of existing spatial structures is rather weak when new industries emerge. Under certain circumstances there are good reasons to assume that place specific features do not determine the location of new sectors. The environment is considered to be of minor importance at the initial stage of development of a sector when there exists a gap between the requirements of the new industry (in terms of knowledge, skills, etc.) and its surrounding environment. Windows of locational opportunity are open in emerging industries.

Thus, during the emergence of a new technology or industry, there will be a relatively large number of locations possessing the generic conditions that would allow the new sectors to grow there. New firms can locate where they please within this variety of places. What then happens, according to this thesis, is that one of these locational contenders is ‘selected’, often by inexplicable chance and random events, so that local enabling effects are very weak (Boschma, 2004; Boschma and Frenken, 2004; Lambooy and Boschma, 2001; Boschma 2004; Krugman, 1991). However, while this interpretation may have some validity when applied to earlier periods of industrial history, the assumption that new paths in the contemporary era require only generic assets that can subsequently be moulded is somewhat questionable. Contemporary innovations combine inherited competences, experiences with customers and many fields of specialized knowledge (Pavitt, 2006). Moreover, this ‘window of locational opportunity’ view at times appears to conflate ex ante unpredictability with ex post inexplicability, and, as we will argue, while some types of selection pressure may be weak during the early phases of industry emergence other types of selection process are critical and strong. As a consequence, this view underestimates the importance of place – of local economic and social context in processes of path creation.

Places influence the origins of new technological-industrial paths in two main ways. First, place is likely to be implicated in processes of entrepreneurial variety generation, and second place is also likely to be important in processes of collective support, selection and the emergence of new trajectories. It is precisely the embedding of agents in particular paths that enables them to
accumulate the resources and experience necessary to launch new technological paths (Garud and Karnøe, 2001). In this sense, the idea of dependence between successive paths, between successive industrial-technological trajectories in particular places becomes relevant. How exactly then, might this occur?

The analysis of the emergence of high technology clusters has emphasised the importance of pre-existing comparative advantages in shaping new increasing returns effects. As Bresnahan et al (2005) argue, ‘old economy’ (or more generally, ‘old’ path) inputs, such as the supply of technical and managerial skills, connections to market niches, and the role of key firms, are often crucial in determining whether and where ‘new economy’ (‘new’ path) based increasing returns actually emerge and develop. While the development of technology clusters can be triggered by specific actions and events with unforeseen effects, these consequences only typically emerge because such events occur in the context of longer term conditions and facilitating factors. Thus while the growth of Silicon Valley was initiated by small contingent decisions such as Shockley’s move to Palo Alto and the founding of Fairchild (Kenney and von Burg, 2001), there were also important preconditions including the development of the electronics industry in the area during the interwar period (Sturgeon, 2001). Another apt example is the Cambridge high-tech cluster (covering scientific instruments, software, and biotechnology, amongst other activities) in the UK. The origins of this are usually ascribed to the establishment of Cambridge Consultants in 1960 (a group of chemistry graduates concerned to foster research links between the University and local industry in the Cambridge area), or to the establishment of a science park by Trinity College in 1970. But they could equally well be traced back to antecedents in the 1930s, such as the existence of Cambridge Instruments (a specialist aeronautical instrument firm) or the local agro-chemical industry.

In fact, there is increasing recognition of the ways in which resources and competences used in old paths may be recombined and reworked to form the basis of purposeful entrepreneurial deviations (Kemp et al, 2001; Metcalfe, 2005; Bathelt, 2001). A key mechanism here is the way in which spin-off firms inherit routines and competences from their parents. There is also mounting evidence that entrepreneurs transfer technological, business and marketing capabilities and competences from parent firms to their new ventures (Klepper and Sleeper, 2005; Buenstorf, 2007). Those spin-offs that inherit routines and competences from parent firms in the same or a related sector have often been
found to be the most successful (Klepper 2007; Buenstorf and Klepper, 2005). Where spin-offs enter a new, but related sector, this may lead to path dependent path creation, as spin-offs increase the diversity of economic knowledge and can evolve in unpredictable ways (Klepper and Sleeper, 2005; Buenstorf and Fornahl, 2006). It is well known that entrepreneurs tend to show locational inertia and prefer not to move home when starting a new venture, partly because of the importance of local social and business communities in mitigating risk (Stam, 2007). Related variety in a regional industrial structure is thus likely to be both an outcome of this transfer, and a resource that facilitates further diversification and innovation (Boschma and Frenken, 2006).

In an important sense, entrepreneurial opportunities are not exogenously given but are deliberately made. New ventures are based on ‘business conceptions’ or the entrepreneur’s interpretation of the opportunity and the approach to exploit it (Witt, 1998; Buenstorf, 2007). Such intuitive conceptions characteristically mix past experience, memory, and current intentions and they have important framing effects on the new firm’s motivation and direction, and its capacity to discover new opportunities. The recognition of technological openings depends heavily on prior experience and learning (Shane, 2000). What is particularly important here is that, as Buenstorf (2007) notes, these conceptions help to pre-select among the variety of potential new activities. *Ex ante* selection is much more significant than implied by accounts that emphasise chance and ‘accidental’ events in path creation, as is illustrated by the example of technological niches.

Technological niches are usually shielded from some market selection pressures as they depend on fringe markets, experimental and special purpose users, or state subsidies and public research funding (Schot and Geels, 2007; Metcalfe et al, 2005). *Ex post* market selection may be weak in these contexts but, at the same time, such niches are typically marked by strong *ex ante* selection in which entrepreneurial agents imagine, conceive, design and gradually improve products and systems that they anticipate will meet demand or provide a basis for new growth. The critical moment here may be the application of existing technology to a new economic domain (Levinthal, 1998). Technological niches are typically marked by unstable technological rules and a lack of an established technological paradigm (Utterback, 1996). Their linkages are friable and fragile. Hence they exhibit a great deal of uncertainty and many information asymmetries and, in such environments, face-to-face contacts and
local networks are important in building trust and knowledge exchange. Given that radical innovation is an inherently uncertain process, there will inevitably be a high failure rate in niche formation. New firms need to identify a fringe market that is not well served by old technologies, products or services, in order to keep themselves alive long enough to develop their own new technologies, products or services (Malerba et al, 2007). The test of a technological niche will depend on whether it can grow in small fringe market niches and then, ultimately, whether it can breakthrough into a more general socio-technical regime (Geels, 2002; 2005; Geels and Schot, 2007). But we suspect that place plays an important role here as technologies, markets and selection pressures co-evolve.

A series of institutional and economic factors at local, regional and national scales, can determine whether an innovation becomes the basis of a new industrial path or whether it remains isolated and underinvested and unable to grow. Local institutions and human resources that have developed as a result of one industry’s development in a region can also be critical causes of, and inputs to, the creation of other industries. These include the entrepreneurial culture; social structures of innovation; access to specialist, demanding and knowledgeable customers; the presence of supporting institutions such as intermediaries, law and venture capital firms; and government provision of hard and soft infrastructures. Carlsson (2007, page 265) for example, writes that “the most important aspect of path dependence may be the existing entrepreneurial climate resulting from pre-existing conditions.” Supporting social infrastructures, however, are difficult to orchestrate in a systematic fashion as they are collective and emergent. Typically, such institutions are not present at the birth of an industry but they gradually evolve as the local industry develops and as processes of positive lock-in consolidate the industrial path and reinforce its momentum (for example, Feldman, 2007). In addition, however, once a supportive generic institutional structure develops for one industry it can have beneficial consequences for subsequent newly emerging paths in other sectors (for example, see Zook, 2005). Thus while supportive local selection environments are typically secondary rather than primary causes of path creation, they can help to develop a technological niche or a radical innovation so that it stands a better chance of surviving market selection pressures.

In summary then, the equilibrium model of path dependence and its geographical offspring accord too little importance to the role of place. In this
The Place of Path Dependence

view, place is confined to a set of large accidental and random initial conditions and triggers, which are followed by a set of cumulative and reinforcing increasing returns. Instead we have argued that place dependence is important well before the unfolding of reinforcing dynamics as it shapes the emergence of paths in particular sites. There is much evidence that local conditions continue to be important to processes of firm spin-off and to the emergence of radically new technological and innovation trajectories. While there is undoubtedly an unpredictable and uncertain dimension to path creation but this should not be exaggerated so as to completely obscure the deliberate ex ante selection of promising entrepreneurial ideas and the creative deployment of pre-existing resources, ideas and relationships. In recent work, chance is conceived as a those random accidents and triggers that occur after the necessary accumulation of antecedent conditions and assets. But this dual-stage perspective can confuse micro-scale events with randomness, and its elaboration will need a much clearer exposition of contingency. Certainly, the place-specific path dependent processes that cross industrial paths are far from deterministic and by no means easy to measure as they will interact in a complex manner with extra-local contingencies and the basic unpredictability of radical innovation. It is not surprising, then, that so far there is very little research that explores these interstitial effects, but it is no exaggeration to say that they are crucial to the long-run adaptability of urban and regional economies.

6. Towards an Ontology of Path Dependent Regional Economic Development

Our aim in this paper has been to examine how far and in what ways path dependence can serve as an explicitly evolutionary concept for studying the economic landscape, thereby focussing on some of the issues that need to be addressed in any attempt to construct what Scott calls an “ontology of regional growth and development that is rooted in the idea of path dependent economic evolution”.

Immediately, a central question has to be confronted: what sort of evolution – of a technology, and industry, or a local economy - is implied by the idea of path dependence? As articulated by its leading exponent, Paul David, the conception is one in which the contingencies of historical accident act to select between multiple (and perhaps competing) possibilities and then various feedback forces come into operation that serve to reproduce the selected
particular technology or industry, a situation that he argues can be thought of as the contingent selection of, and lock-in to, one of a number of possible multiple equilibrium states. In his model it takes an exogenous shock to disrupt that equilibrium state, so that economic evolution takes the form of successive punctuated equilibria. David refers to this conception as ‘strong history’ (see also Castaldi and Dosi, 2006).

We have argued, however, that this conception is overly restrictive, too ‘strong’ we might say, and that David's very notion of 'path dependent equilibrium economics' is itself something of a contradiction in terms. Of course there are examples of this sort of selection and lock-in of a technology and even an industry to a stable, self-reproducing state; but even such instances need not represent equilibrium situations. According to complexity theory, for example, it is possible for a system to exhibit stability or inertia even though it is ‘far from equilibrium’. Moreover, examples of technologies, industries – let alone whole local or regional economies - being in a stable and self-reproducing unchanging state are not the norm. Most technologies, industries - and local and regional economies – follow development paths that evolve over time, in a path dependent manner. The idea of multiple equilibria, to our mind, fails to capture this process. We have sought to argue, therefore, for a richer interpretation of path dependent economic evolution, one that does not require or necessitate notions of equilibrium.

Instead, we conceive of the idea of path dependence as entirely consistent with patterns of economic evolution in which technologies, industries, institutions, and regional economies adapt and mutate over time without ever reaching or tending towards any equilibrium. David, no doubt, would criticise this view as one of ‘weak history’. But, arguably, it is a history that accords with much of observable reality. Whereas in David’s view of path dependence, history to all intents and purposes actually ceases once the system in question becomes locked into it equilibrium state, in our conception history continues to unfold, in a path dependent manner. This means that the focus of analysis centres on how economic evolution takes place along paths, and on whether and in what ways such paths undergo ‘life cycles’ (and thus how processes of ‘positive’ path dependence turn into or are replaced by processes of ‘negative’ path dependence), and on how old paths are replaced by new.

The sort of ontology we have in mind has similarities with the way that some sociologists interpret the idea of path dependence in terms of ‘reactive sequences’ (Mahoney, 2000, 2006). Reactive sequences are chains of temporally
ordered and causally ordered events. In a reactive sequence, each event in the historical sequence is both a reaction to antecedent events and a causal influence on subsequent events. This still means that early, initial events are important to later outcomes: small changes in initial conditions can accumulate over time and make a great deal of difference by the end of a sequence. In this respect, reactive sequences retain one of the key aspects of the path dependence idea. But in other respects, reactive sequences are quite different from the basic ‘lock-in to equilibrium’ conception of David-type path dependence models. Whereas the increasing returns sequences that are typically invoked in standard path dependence models are characterised by processes of reproduction that reinforce early events, reactive sequences are marked by processes that transform earlier events: in a reactive sequence, initial events trigger subsequent development not by reproducing a given pattern, but by setting in motion a chain of linked reactions and events. As such, this interpretation of path dependence offers a much more open set of possibilities in terms of the trajectories of technological, industrial and regional economic paths. If the objective of a path dependence approach to understanding regional growth and development is indeed to uncover and make sense of the specific historical mechanisms and chains of events that have produced a particular economic landscape (rather than some other possible configuration), then the idea of path dependence as reactive sequences would seem to be well worth exploring in evolutionary economic-geographic work.
References


and Management, Sant’Anna School of Advanced Studies, University of Pisa.


A stochastic system is called ergodic if it tends in probability to a limiting form that is independent of the initial conditions. Breakdown of ergodicity gives rise to path dependence. When path dependence occurs, ‘history matters’. Choices made on the basis of transitory conditions can persist long after those conditions change. Path-dependent features of economics range from small-scale technical standards to large-scale institutions. Prominent path-dependent features in economics include technical standards, such as the 'QWERTY' standard typewriter keyboard and the 'standard gauge' of railway track. Ergodicity and breakdown thereof plays a major role in models of social interaction.

What we have in mind is a path dependence model of the sort shown in Figure 2. New paths do not emerge in a vacuum, but always in the context of existing structures and paths of technology, industry and institutional arrangements. These existing structures and paths - that together constitute the ‘pre formation phase’ - provide the stimulus for, and shape the scope of new opportunities, of technologies, products, whole industries, and institutions. Several different alternative new technologies or products may co-exist at this stage. Which particular technology or product or industry out of the alternatives emerges - or is ‘selected’ – may be simply be a chance or contingent event, for example where conditions happen to favour one alternative rather than another, or where one alternative has a slight ‘first-mover’ advantage over others; but could also be the result of deliberate and purposive (and competitive) behaviour by economic agents, the birth of a new technological, industrial or institutional path from
possible alternatives – the path creation phase - may be the result of accidental or happenstance factors, or the result of deliberate and purposive behaviour by economic agents.
Figure 2. A Revised View of the Evolution of an Industry or Technology as an Evolutionary Process