**THE EMPIRICAL INVESTIGATION OF NON-LINEAR DYNAMICS IN THE SOCIAL WORLD – ONTOLOGY, METHODOLOGY AND DATA [1] 22/02/19**

**Graham Room, University of Bath**

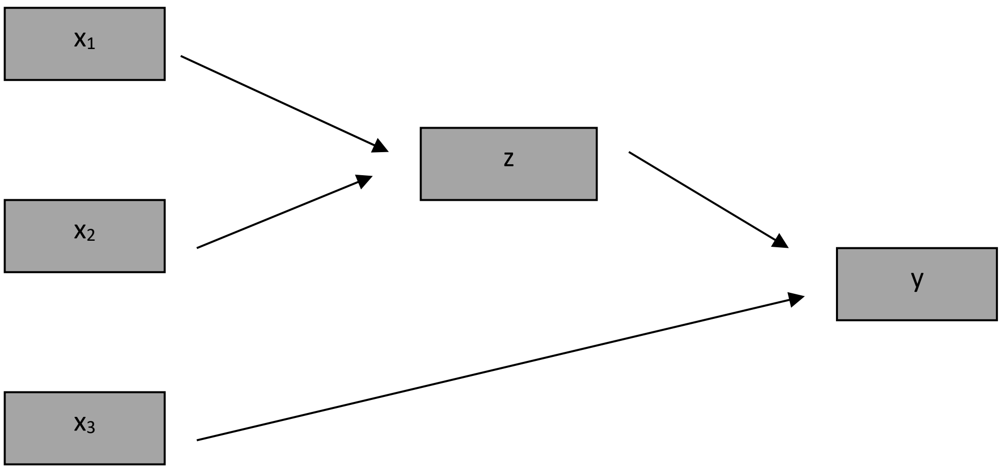
**ABSTRACT**

Across much of social science, quantitative methods involving linear regression reign supreme. This article makes the case for studying co-evolving systems, as a form of *non-linear* dynamics. If such investigations are to be fruitful, they must not only be elaborated theoretically, they must also be applied to empirical datasets. This article considers how this can be done, with what sorts of data sets and what forms of data analysis. It takes as its specific example the international datasets on patents, as revealing processes and patterns of technological innovation. It shows how such an approach can illuminate scholarly debates and develop indicators for policy makers. Finally, it offers an agenda for research into dynamic co-evolving systems across other empirical areas.

## 1. INTRODUCTION - THE GENERAL LINEAR MODEL

The task of social science is to explain social phenomena. This, it is commonly asserted, should involve measuring the effects of different 'independent' variables on some 'dependent' variable of interest. This can be presented diagrammatically. In Figure 1, the independent variables $x_1$, $x_2$ and $x_3$ shape the dependent variable $y$ (but with some effects exerted via the intermediate variable $z$).

**Figure 1: The General Linear Model**




This vision or ontology of the social world can be presented as a set of equations:

$$y = f_1(x_1) + f_2(x_2) + \dots f_n(x_n) \qquad (1)$$

---

[1] This paper builds on two earlier programmes of work by the author:

- An FP7 project for Eurostat in 2001-4: New Economy Statistical information System (NESIS), on statistical indicators of the European knowledge economy, directed by Deo Ramprakash
- An ESRC Fellowship in 2008-10: Agile Actors on Complex Terrains (Award RES-063-27-0130).

It also benefits from the work of the DCICSS (Dynamics of Cumulative Innovation in Complex Social Systems) group based at the University of Bath, which includes Evangelou Evangelis, Orietta Morsili, Lorenzo Napolitano, Emanuele Pugliesi, Alastair Spence, Paolo Zeppini. Francois Lafond gave valuable advice on the patent classification process.

Quantitative social science applies this vision using regression analysis. It commonly casts the problem of explanation in terms of a set of linear equations; this is why it is often described as the 'General Linear Model' (Abbott, 2001: Ch 1). The GLM looks for the straight line that best estimates - and therefore 'explains' - the dependent variable *y* as the additive outcome of a number of independent variables $x_1 \ldots x_n$ plus a random error term *u*:

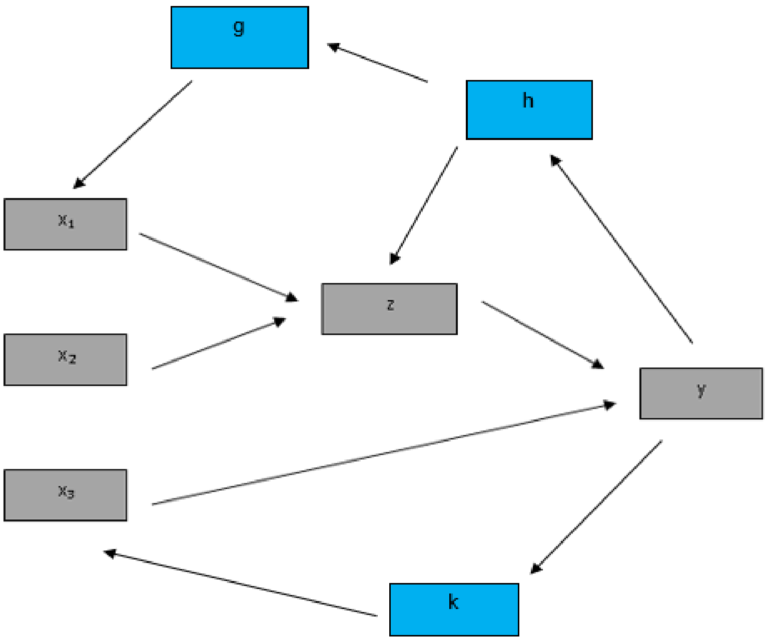
$$y \quad = \quad b_1 . x_1 + b_2 .x_2 + \ldots . \; b_n .x_n + u \qquad (2)$$

The *b* coefficients measure the rate at which changes in the independent variables *x* produce changes in *y*. The error term *u* is a measure of how closely our equation captures the empirical data – and how much 'noise' remains around its predicted values of *y*.

There are a number of assumptions involved here however, which should be made explicit.

1. The GLM assumes that the separate causal effects of the independent variables can be isolated – there are no significant interactions among them. When interactions are substantial, it may be possible to reduce them by re-defining the independent variables.

**Figure 2: Feedback processes**




2. The GLM assumes that the influence of the dependent variables *x* on the dependent *y* is downstream and one-way. There are no significant feedback processes allowing *y* to influence *x*. And yet of course, in the real world such feedback processes are common. This may mean that instead of a downstream and uni-directional determination of *y* by *x*, the directions of influence run upstream as well (**Figure 2**). Such self-reinforcing processes are well-recognised across the social sciences, including for example the economics of 'cumulative causation' in Myrdal and Kaldor, the system dynamics of Checkland and Coyle, and the social policy literature on the dynamics of social exclusion (Toner, 1999; Powell and Bradford, 1998, 2000; Room, 1995).

Nevertheless, the mathematical modelling of such cumulative feedback is under-developed across much of social science. [2]

3. The GLM assumes that changes in the independent variables will produce proportional changes in the dependent variable, across the entire range of observable values, within a rather simple 'timescape'. In the presence of the afore-mentioned feedback processes however, timescapes are more complex. Some effects of the independent variables are short-term and immediate, while others are long-term and delayed (Abbott, 2001: Ch 1). Many social processes are replete with time lags, ratchets and path dependencies (Lieberson, 1987: Ch 4). Pierson (2004) points to the consequences of long-term and often slow changes in background social and economic conditions (pp 74-7). There may be long periods of stasis, and then thresholds at which sudden avalanches of reconfiguration occur, sometimes discussed in terms of 'punctuated equilibria'. It is possible for the GLM to handle some of these temporal complexities, using more sophisticated statistical techniques. Nevertheless, taken together, they show the GLM struggling, when applied to social sub-systems and processes with complex inter-connections.
4. The GLM assumes, finally, that the independent and dependent variables are given and that the relationship among them is fixed. This may well be appropriate, at least in the short term. 3 Sooner or later however, the relationships among the 'variables' may change and the variables themselves amalgamate, divide or disappear and new ones emerge. For understanding these dynamics, the techniques of the GLM are of little use.

It is hardly surprising that the GLM exerts such a powerful sway across many disciplines. It is readily visualised; it can be formalised in terms of simple equations; it is convenient and tractable; it can often be applied to good effect. At the very least, it is often a useful first step (Jervis, 1997: 260-61). Nevertheless, where the foregoing assumptions fail to hold, its general applicability to the real world is put in question and we must consider alternatives.

## 2. CO-EVOLUTIONARY DYNAMICS

The GLM finds its inspiration in classical physics. Evolutionary biology provides a quite different source of conceptual and methodological inspiration – albeit one which social scientists have interpreted and applied in a diversity of ways. [4]

The strengths of an evolutionary model align closely with the limitations of the GLM discussed in the previous section. With evolutionary science as our starting point, we now therefore consider co-evolutionary dynamics, as an alternative to the GLM.

In his account of the diversification of species, Darwin was centrally concerned with processes of adaptation to different habitats. He depicted this visually as a 'Tree of Life' (Darwin, 1859). This offered successively sprouting branches and sub-branches, as particular 'variations' adapted to and exploited different habitats, over many millions of years.

Imagine looking down on the tree of life from above and viewing the top-most branches. Displayed there are the various species that are alive today (or we could consider any other horizontal 'cut',

---

[2] It is common to use lagged variables to build in some dynamics, with $y_{t-1}$ being used as a predictor of $y_t$. This does not however allow for the feedback processes from *y* to *x* that Figure 2 displays.

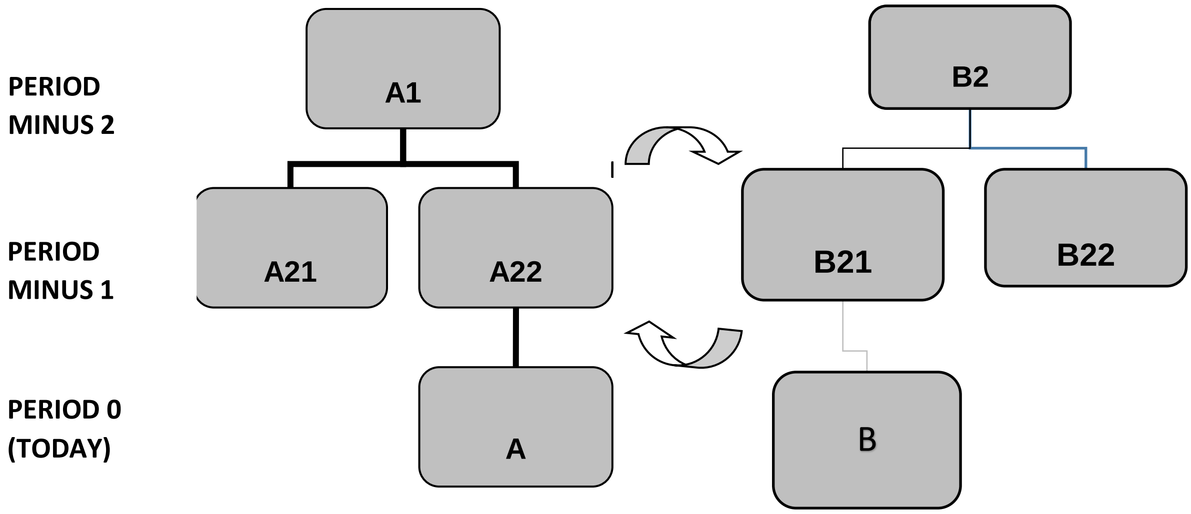
[3] It was for example the basis for Alfred Marshall's treatment of the short-term: the stock of capital in an economy was given, as was the relation between different factors of production. In the long-term however, these were all malleable.

[4] For an eloquent statement of the relevance of evolutionary models in social science, closely consistent with the argument of the present chapter, see Lieberson and Lynn (2002). See also Hodgson (2001: Ch 22; 2002); Blaug et al (2011). Contrast this with those who retain a focus on the biological substrate of human behaviour (for example, Sloan Wilson, 2008).

representing the species that lived at another chosen period in the Earth's history). Across each of these cross-sections, the various branches (species) are connected in ecosystems of interdependence, involving dynamic synergies and arms races of co-evolution (Kauffman, 1993; Maynard Smith and Szathmary, 2000). These typically involve populations far removed from each other across the evolutionary tree: for example, flowers and insects. They powerfully influence which species thrive and which are extinguished.

Figure 1 provided a visual representation of the GLM in its most basic terms. Such diagrams can provide powerful images that organise and direct our thinking about a given phenomenon. We now therefore consider (**Figure 3)** a corresponding representation of evolutionary dynamics. This brings centre-stage variables that emerge, divide and disappear, as their interrelationships unfold. In subsequent sections we apply this to the social world.

**Figure 3 Co-evolutionary Dynamics**




**A** and **B** (at the bottom of the diagram) are two mutually adapted entities in the world of today. They might for example be the populations of two species such as bees and flowering plants, each benefitting as the other thrives.

We then pose the question: how did this mutual adaptation arise? We decline to treat it as a causal correlation, with the population of flowers 'causing' the population of bees within a timeless environment. Instead, we seek to unpick the intricate and messy history of successive contingencies that has led to the mutual adaptations of today.

The upper part of the diagram reveals those historical contingencies. **A1** and **B1** were the ancestors of today's bees and flowers. As we know from Darwin, in each generation, variations are produced. In general however, as long as the environment remains stable (**Period -2**) they are unlikely to displace **A1** and **B1**.

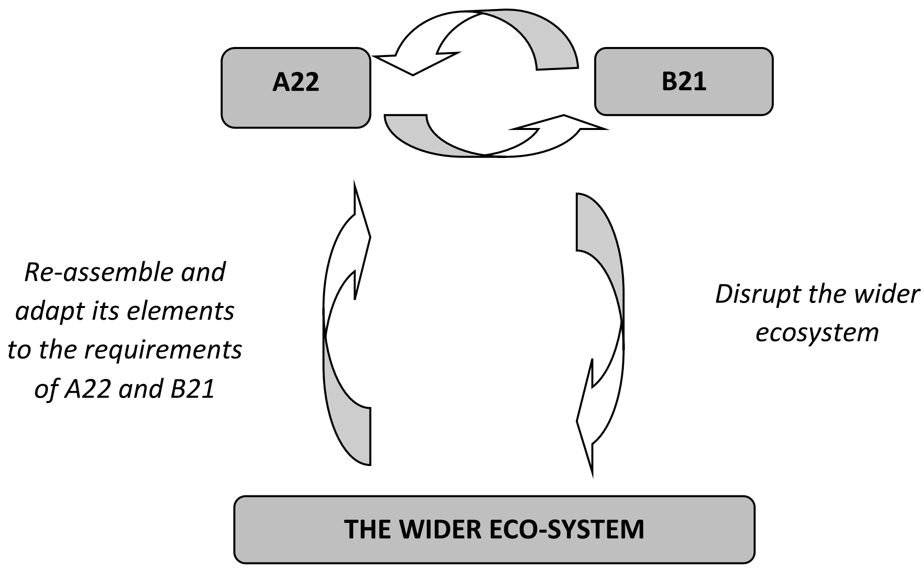
It is when some environmental change occurs, at the start of **Period -1**, that we may expect some of the variations (**A21, A22, B21, B22**) to be adopted as superior to **A1** and **B1**, which now become extinct. However, which of the variations **A21** and **A22** becomes preponderant depends in part on the new biotic environment constituted by the arrival of **B21** and **B22**; and *vice versa*. In short, what is crucial is which of the four sets of interactions between **A21** and **A22** on the one hand, **B21** and **B22** on the other, is of greatest mutual benefit.

In the diagram, we show the relationship of **A22** with **B21** as being this favoured pairing, this synergy or 'elective affinity'.[5] Each will now accelerate the flourishing of the other. Their flourishing will in turn deny resources to **A21** and **B22**, which in the 'struggle for existence' become extinct. Hence we arrive at the bottom row of the diagram, **Period 0**, where **A** and **B** dominate. Here, by virtue of their domination, the environment is quite different from that in **Period -2** or even in **Period -1**. And indeed, **A** and **B** are themselves quite different in their capacities from their respective forebears **A1** and **B1:** perhaps barely recognisable. Nevertheless, the domination of **A** and **B** is unlikely to last for ever; further rounds of interaction with the wider eco-system will eventually destabilise it, as new rounds of variation and selection are set in motion.

The dynamic synergies among particular elements have, as their obverse and corollary, the progressive disruption of other connections and elements of the eco-system – and the incorporation of those elements into the 'empire' of the favoured elements (see **Figure 4**). The **A22-B21** axis becomes a vector of cumulative change, around which the wider eco-system is progressively re-ordered and re-configured. This also makes it a non-linear system with strong path dependency, where instead of the additive effects that are central to the GLM, change is multiplicative and self-reinforcing.

**Figure 4 Creative Disruption**

**The dynamic synergies between A22 and B21 disrupt and re-cycle the wider eco-system**




[5] The term 'elective affinity' was originally used in German chemistry of the 18th Century, to refer to the way in which compounds interact and combine selectively with each other (Howe, 1978). The search for such affinities in chemistry was conducted in the shadow of Newton and in envy of physics and its claim to universal natural laws - just as our own account has been located within the larger debate about social science and the Newtonian legacy of the GLM. Goethe took this idea of elective affinities into his novel *Die Wahlverwandtschaften,* applying it to sexual attraction. Kant in turn applied the idea to relationships among concepts; Weber to relationships between ideas and the interests of social actors. Perhaps surprisingly however, it does not seem to have been used in relation to biological co-evolution. In these various cases, 'elective affinity' is not just a matter of complementarity or similarity; it is a dynamic synergy, in which elements that are especially favourable to each other enable the ensemble as a whole to flourish. It thus offers a dynamic of mutual selection, reinforcement and change. Crouch (2005: Ch 3) has been a trailblazer in applying such a perspective within institutional sociology and anticipates much of what is said here.

The dynamic synergy cannot however continue without limit: nor can the concomitant disruption and recycling of other elements. Some parts of the wider eco-system are too resilient and robust to be unpicked and re-worked: they constitute an 'evolutionarily stable state' (Maynard Smith, 1982). Moreover, the cumulative change that is driven by the elective affinity of **A22-B21** is forever opening up new possibilities - even while **A22** and **B21** flourish more than ever before, new windows of opportunity have thereby been opened for other elements of the ecosystem, other dynamic synergies that may eventually match or even surpass **A22-B21**.

## 3. THE MATHS OF NON-LINEAR CO-EVOLVING SYSTEMS

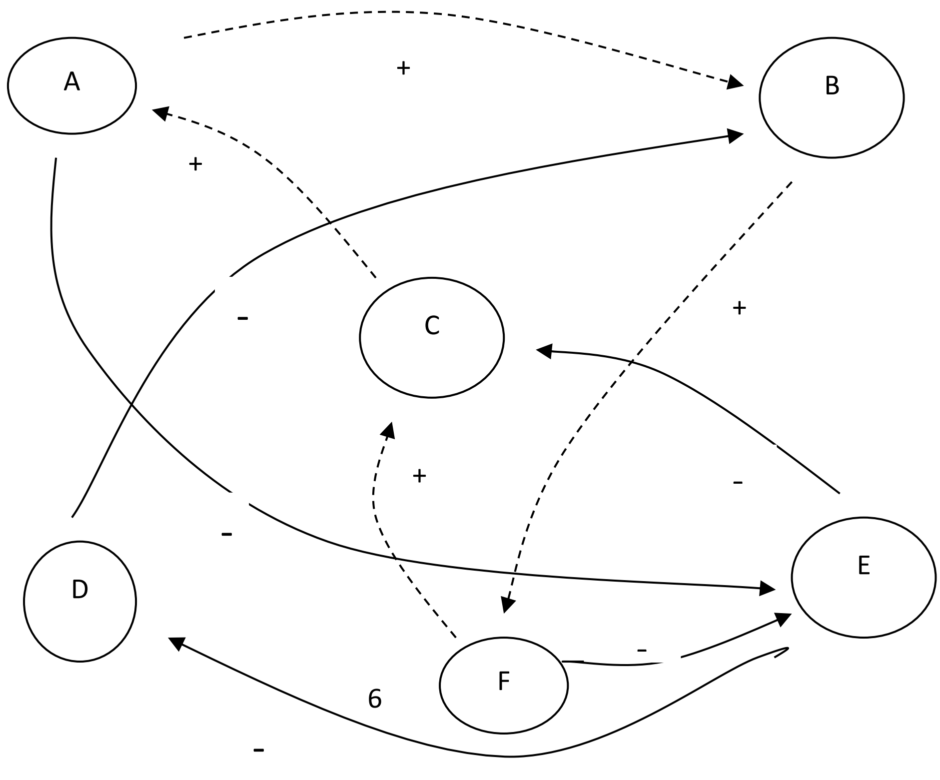
We now require an empirical methodology, appropriate to the dynamics of co-evolving systems and applicable to the social world. In developing this methodology, we again take inspiration where appropriate from evolutionary biology, as well from social scientists who have sought to capture these dynamic processes in their empirical research.

At the heart of co-evolutionary dynamics – and their parsimonious depiction in Figure 3 – is a dynamic ensemble of mutually adapted elements. This is an ensemble that has emerged from processes of mutual reinforcement: a process that crowds out or progressively dominates other elements. We seek an empirical methodology for analysing such dynamic ensembles in the social and political world and the processes of mutual adaptation from which they have emerged.

***Qualitative System Dynamics***

Powell has developed qualitative system dynamics (QSD) for the analysis of institutional change; he builds on the work of such writers as Checkland and Coyle (Checkland and Scholes, 1990; Coyle, 1996; Powell and Bradford, 1998; Powell and Bradford, 2000). He first maps the organisations of interest and the connections of interdependence they involve. He then labels each line of interdependence, to indicate its direction but also whether the relationship is direct or inverse - whether, in other words, an increase in some property or activity of the 'upstream' node causes a change in the 'downstream' node that is positive or negative.

**Figure 5 Runaway Loops**



Within this map, Powell proceeds to identify those cycles whose links are all positive. These are cycles which loop back on themselves in self-reinforcing circles. When any one element starts increasing, the whole sub-system experiences explosive growth; when any starts decreasing, the sub-system experiences implosive collapse. Powell refers to these as 'runaway loops'. In **Figure 5** the runaway loop is marked as a dotted line. This is the counterpart to the elective affinity between **A22** and **B21** in Figure 3. Meanwhile other cycles of interdependence loop back on themselves, in ways that dampen down change and stabilise the system as it presently exists.

Having identified the runaway loops, the next step is to assess how strong is the self-reinforcing dynamic. This will determine the speed at which the cycle 'runs away'. Second, we need to know how well-connected the sub-system or loop is to the system as a whole, so that its runaway loops have wider influence. There may be particular threshold effects: beyond a certain point, the runaway sub-system triggers other sub-systems, which in their turn begin also to 'run away'. The particular dynamics that emerge will be heavily dependent on the way that elements are connected to each other.

Nevertheless, Powell's QSD assumes that the configuration of elements – in Figure 5 for example – remains fixed. There are no novelties and no extinctions. He does not allow for the sub-system that runs away to change its configuration or to re-shape the larger system in which it is embedded. Powell advances our quest: but he does not provide us with a co-evolutionary dynamic.

***Autocatalytic Sets***

Jain and Krishna (2003) are interested in autocatalytic sets (ACS) and the key role that these appear to have played in the origins of life. An ACS comprises a set of simple molecular organisms, none able individually to self-replicate, but each providing a catalyst for each of its fellows - a process of symbiotic and co-evolutionary 'boot-strapping' for collective self-replication (see also Kauffman, 1993: Ch 7). This is of central relevance to any Darwinian account of the origin of species.

Like Powell, Jain and Krishna employ a methodology of directed graphs (networks where the *direction* of the connection matters). [6] They model a population located at each of the different nodes, dependent on the growth of population at a number of other nodes. Their first step is to watch which nodes thrive and which do not – and the significance of the ACSs for this variation in fortune. They notice (Section 3) that the population of a node enjoys particularly rapid growth if it is part of an ACS. An ACS thus plays the same role in Jain and Krishna as 'runaway loops' play in Powell.

So far, this analysis by Jain and Krishna - like that of Powell - does not allow for any reconfiguration of the system or any transformation of its elements. This we will refer to as the '*fast' dynamic'*, as the nodes within a given configuration of the network progressively populate the various nodes, eventually reaching an equilibrium.

Jain and Krishna then however apply a Darwinian rule, to effect a re-configuration of the network (Sections 4-5). This we will refer to as the '*slow' dynamic*, occurring only once the fast dynamic has produced an equilibrium. The nodes that flourish least are extinguished; random new nodes and connections are then added, mimicking the mutations in Darwin. Such reconfigurations are of course central to any evolutionary framework. This formal analytic thus captures co-evolutionary dynamics in elegant and parsimonious fashion.

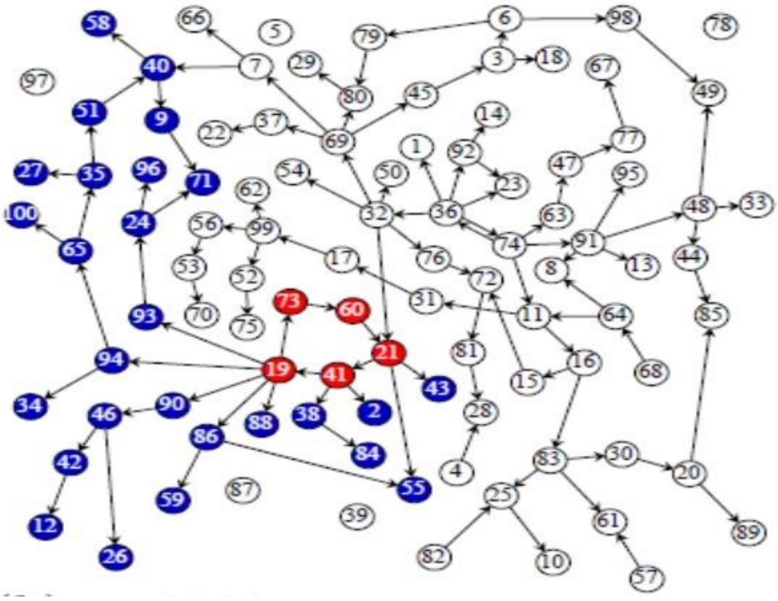
Jain and Krishna use computational models to simulate what can happen in such a dynamic system, depending on the parameters and algorithms adopted. For one moment in these successive transformations, see **Figure 6**. One ACS has developed, centred on the five nodes coloured red, but also benefitting a range of blue nodes, which thrive on their accelerated development, while not themselves contributing to the self-reinforcing feedback loops. The white nodes however are left

[6] Also available at http://arxiv.org/PS_cache/nlin/pdf/0210/0210070_v1.pdf

without benefit from these elective affinities; they will therefore provide the low fitness candidates for extinction.

Computational models of this sort can then serve as ideal types, illuminating the range of empirical dynamics to be found in the real world (Gilbert and Troitzsch, 2005). To do this is likely to be rather demanding in terms of data, depending as it does on details of connections across a complex system, collected repeatedly and in timely fashion, especially if the aim is to engage with policy audiences. This is an obvious example of 'big data' analysis - on the one hand drawing on the large-scale administrative data that are routinely amassed and regularly updated; on the other hand, deploying modern computational capacity to scan such data sets, for the 'runaway loops' and the ACSs they reveal (Mayer-Schonberger and Cukier, 2013).

**Figure 6 Jain and Krishna's Autocatalytic Sets**



*Red nodes belong to the core of the dominant autocatalytic set of the graph, blue nodes to its periphery, and white nodes are outside the dominant autocatalytic set. This diagram shows run 6062 in the computational modelling of the graph (Jain and Krishna, 2003).*

***Eigenvalues and Eigenvectors***

A linear system tells a story of one-way influence or determination. Rising unemployment causes growing poverty; rising obesity causes higher rates of diabetes. The independent variables exert a force on the dependent variable - Newton's mechanics applied to the social world.

In a linear system, as captured in Figure 1, the *x* variables are independent both of each other and of the dependent variable. What however if there are feedback processes, of the sort depicted in Figure 2, including a variety of 'cycles' (loops), such as ***z-y-h*** and ***$x_3$-y-k***? These cycles allow self-reinforcing forces to develop, similar to those highlighted by Powell and by Jain and Krishna. The influence of the independent variables on the dependent variable is no longer one-way; and the independent variables are no longer isolated from each other. Here instead is a network of interacting nodes, with the rate of activity on each node determined by the activity levels on the other nodes to which it is connected.

Mathematically, we can present the connections of a network as a matrix (the 'adjacency matrix' of the network) - with as many cells, both vertically and horizontally, as there are nodes. The cells of the matrix show whether any two nodes are connected and, if so, in which direction. This could, for example, be done for the networks in any of the Figures 2, 5 and 6.

In equation (1), the function $f_i(x_i)$ showed how $y$ derives from $x_i$. In equation (3) the matrix **C** plays a somewhat analogous role, where the coefficients of **C** relate **da/dt** (the change in the activity level $a_i$ on node $i$) to the activity levels on other nodes. It thus allows us to see how the activity levels on other nodes drive the change in activity level on whichever node is of interest.

$$\mathbf{da/dt} = \mathbf{C \cdot a} \quad (3)$$

In equation (3) the activity levels $a_i$ on node $i$ are collected in the vector $\mathbf{a} = [a_1 \ldots a_n]$.

This matrix **C** of interactions can be expressed more parsimoniously, in terms of its **eigenvalues** and the corresponding **eigenvectors**. The eigenvalues measure how intense are the interactions among the nodes; the eigenvectors partition the matrix into corresponding zones. Together they then allow us to summarise the evolution of the dynamic system in equation (3) with elegant simplicity. The eigenvalues display the strength of the self-reinforcing forces unleashed by the interactions among nodes; the eigenvectors show where in the system they act.

A matrix such as this may have a number of different eigenvalues and corresponding eigenvectors. It is the eigenvalue of largest (absolute) value that dominates the dynamics of the system as a whole. By identifying the largest or most powerful (real) eigenvalue in the system and the direction of its eigenvector, we can identify the final configuration of activity levels it will produce (its 'attractor' or equilibrium). [7]

Jain and Krishna attach major significance to the eigenvalues of the network. The larger the eigenvalues, the stronger the ACSs to which they correspond and the more pronounced the basins of attraction they define. In Section 6 of their paper, these eigenvalues play a key role in their analysis of the various types of innovation that develop. In applying their approach to our empirical case studies, the eigenvalues and eigenvectors will likewise be our key tool for making sense of the ACSs that develop and their likely dynamics.

Eigenvalues can be positive, negative or complex. The resulting dynamics are sharply different, whether accelerating, dampening or spiralling. In many engineering systems, the strength and direction of these forces can be a major challenge - but also an opportunity, to be monitored and mobilised. Our aim in this paper is to demonstrate the utility of such analytical methods for social scientists, but also their practical value for policy actors.

## 4. CONTINGENCY AND THE ARTS OF CIVILISATION

In the GLM, the variables and their boundaries are fixed. So are their interrelationships. The practitioner of GLM seeks to measure the effects of the independent variables on the dependent variable as a matter of linear addition. Much energy has been devoted to clever ways of doing this, even under apparently unpropitious circumstances.

Co-evolutionary dynamics involve a quite different notion of causation. Here are phenomena whose configuration and frequency are driven by a dynamic that is synergistic not additive. It is the 'elective affinities' among particular elements that lead to their progressive domination of the system in question and the corresponding reconfiguration of the variables.

[7] There is a parallel to be drawn with the GLM and equations (1) and (2). There we might estimate the importance of the different ***x*** variables in predicting the value of ***y*** and rank them by reference to the proportion of variance they explain. We might then, in a spirit of parsimony, include lower-ranked x variables only up to the point where we have explained our desired proportion of variance.

The fruit of any such enquiry is not however a set of 'universal truths', but instead some 'time-bounded truths' about the contingent dynamics of change (Brown and Langer, 2011). Thus when Jain and Krishna run successive simulations of their model, small shifts in the algorithms of extinction and mutation can produce quite different trajectories. This is consistent with Gould's (1991) argument that if the biological 'tape of history' were to be 'replayed' many times, altering some apparently minor detail in that chain of contingencies, each would have produced a quite different result.

Few of those replays would have produced anything like *homo sapiens*. That however made all the difference. *Homo sapiens* now purposefully intervenes in the contingent dynamics of co-evolutionary change, shifting the algorithms of extinction and mutation and the terms on which the tape of history is played. Thus in some degree, our species makes its own history, not least in regards to the biological world of which Darwin wrote and where human interventions are producing new mass extinctions. It is imperative therefore that we consider how to conceptualise purposeful agency as exercised by human actors. There are three aspects to consider (Room, 2012).

***The driver of evolution: artificial or natural selection?***

Darwin was concerned not only with the blind dynamics of natural selection, but also with the purposeful interventions that human beings undertake, to modify the results of those dynamics. His *Origin of Species* began with the practises of the pigeon breeders and horticulturalists whom he knew (Darwin, 1859: Ch 1). They looked for novel characteristics in the offspring of each new generation, which would better suited to their requirements as breeders. These offspring they then bred for future use. Darwin saw that in nature, but over a much longer period, the struggle for existence would play an analogous role but without purposeful agency.

Darwin thus made the mental leap from 'artificial selection' to 'natural selection'. In applying his model of evolution to human societies, we re-engage with the practices of active husbandry from which Darwin began. These are the arts of civilisation. Instead of the blind adaptation of species to a harsh environment, they involve human reflection, experimentation, collaboration and the growth and application of knowledge. Nevertheless, human beings still live within a natural world, variously intervening and modifying the processes of natural selection, both thoughtfully and thoughtlessly. The 'natural' world we see around us is the result of both natural selection and such acts of purposeful intervention - natural and artificial selection in an endless dance.

***The locus of evolution: biology or technologies?***

Those who champion evolutionary ideas and want to apply them to the social world disagree as to *what it is that evolves*. For Dawkins, it is a matter of understanding social dynamics by reference to the demands of biological evolution. It is for example by reference to the 'selfish gene', that we may understand the evolution of cooperation and altruism (Dawkins, 1976). Evolutionary economists in contrast leave no place for biological selection (Hodgson, Potts).[8] New 'variations' emerge from the 'animal spirits' and inventiveness of entrepreneurs, in what Schumpeter described as 'swarms of innovation'. In Darwinism, the genetic legacy of a species is re-worked; in evolutionary economics, it is the technological and institutional legacy of a society (Potts, 2000; Beinhocker, 2007; Crouch, ). That includes the technologies and institutions by which we manage the 'natural' world and extract benefit from it.

Just as pigeon breeders and horticulturalists looked out for novel characteristics in the offspring of each new generation, entrepreneurs are forever on the lookout for new technologies whose co-evolutionary dynamics could open new markets and yield disproportionate returns. Such dynamics

[8] This does not mean overlooking that human beings are biological organisms. They feed on other organisms; they are vulnerable to the ravages of new viruses; much of their economic and social activity is geared to the collective management of these challenges (Flannery, 1994: Part 2). Nevertheless, the variations thrown up in their social and economic technologies - and then variously selected and retained - are not biological. It is in this narrow but crucial sense that the analysis of societal evolution can and should ignore the biology.

may entail co-evolution between different technologies; between new technologies and new markets; between new forms of industrial organisation and new systems of public regulation, etc. To discover and nurture such dynamics is central to what we are calling the arts of civilisation (Bronowski, 1981: Chs 2-4). [9] Those dynamics are in some degree 'blind' and their outcomes unpredictable, depending on how they fare in the market place; nevertheless, they are also subject to purposeful interventions, including by large corporations and public policy makers. This too is an unending dance.

***Power, purpose and positional advantage***

How a given technological innovation will fare – and how it may interact with other technologies and institutions – can never be entirely foreseen. Entrepreneurial ingenuity may *propose* new variants: but it is processes of differential selection through the market that *dispose*; and these can seem just as collectively 'blind' and devoid of overall intent as the processes of natural selection that drive speciation in the wild.

Even as those collectively blind processes unfold however, strategically purposeful social actors will attempt to modify them to their own ends – depending on the resources and power at their disposal. This is a struggle for positional advantage. Attempts to apply evolutionary models to the social world have in general neglected such exercise of power. We bring power centre-stage: all set within the hierarchical relationships of domination and dependence, which characterise our human societies.

We will henceforth speak of the **Contingent Historical Model** (CHM) as our alternative to the General Linear Model: integrating co-evolutionary dynamics with purposeful agency and political economy (Room, 2011, 2016). It is this model that we seek to apply empirically and in ways that engage with both scholarly and policy concerns.

## 5. THE INVESTIGATION OF EMPIRICAL DYNAMICS

For phenomena that involve co-evolving systems, the GLM is inappropriate. We have argued instead for the **Contingent Historical Model** (CHM). In doing so, we will need to take into account the unending 'dance' between artificial and natural selection, applied to the world of technologies and institutions.

We will use the Jain and Krishna model of co-evolutionary dynamics as the basis for our empirical social research. In the present paper, the focus is on new technologies, as captured in patent databases. We will show how the Jain and Krishna model can be applied to patents and what empirical analysis is possible. Depending on their distinctive features, for future case studies we may need to modify the model; and there are always likely to be trade-offs between the realism of the model and its mathematical tractability.

We construe the world as a connected network of loci or nodes. The level of activity on each node is affected by that at the other nodes to which it is connected. This is the '**fast dynamic**' of Jain and Krishna. The pattern of outcomes in such a vast connected system cannot however be predicted in advance, as the simple aggregation of activity at the micro-level - it is blind and emergent and the outcome may be counter-intuitive. This is a general feature of complex systems: see for example Schelling, 1978; Squazzoni, 2012).

The **'slow dynamic'** is different. It involves the extinction of nodes with low levels of activity and the introduction of new nodes and connections. In Jain and Krishna, these are random novelties. It can also however be effected by purposeful selection by social actors. This may happen at the local level, with inventors producing new technological devices (Koenig et al., 2008), a few of which find elective

[9] This is very much in the tradition of economics writing on 'cumulative causation' including in particular Kaldor (Toner, 1999). It contrasts markedly with orthodox economics and its preoccupation with market 'equilibrium' (Kaldor, 1972). Also relevant here is the literature on national innovation systems (Lundvall et al., 2006).

affinities that enable them to thrive disproportionately. It can happen on a larger scale, when the big actors of government and the corporate world consider the emergent macro-outcomes of the fast dynamic and purposefully intervene, hoping to steer them in new directions. It can also involve meso-level actors mobilising from below, to capture those global dynamics and impose agendas of their own (Kristensen and Zeitlin, 2008; Goodman and Thornton, 2018).

There is a large and diverse literature that touches on this sort of 'slow' intervention within complex and emergent dynamics. We take two quite different examples (more generally, see Room 2011, 2015; and Kingdon, 1995).

Stewart (1997: 96-7, 317ff) is concerned with the stability of complex engineering systems (the dynamic counterpart to a thermostat in a central heating system.) He studies how a system's trajectory can be steered by repeatedly adjusting the system parameters. This will involve discontinuous interventions - the seizing of critical moments, the throwing of particular switches - rather than a continuous and smooth process. [10] Concerned as he is however with engineering systems, Stewart ignores the hierarchical relationships of domination and dependence, with which those studying social systems are unavoidably concerned, and the significance of power, political economy and the struggle for positional advantage.

Those latter factors are, in contrast, central to Freedman's (2013: Ch 33) study of strategy, both in business and in war. He contrasts actions that express the playing out interests *within* a given hierarchy of power and social organisation; and properly *strategic* actions aimed at *changing* that organisation. This is a purposeful slow dynamic: involving, as it does, strategic imagination of the new dynamics that specific interventions are expected to set in motion; a sense of the critical moments when those interventions will have the desired effect; and an ability to read the 'weak signals' of impending change.

In both examples, the slow interventions involve 'seizing the moment', when the system reaches a certain critical state, and 'reading' the system dynamics and the alternative directions in which it might now move. As seen above, it is the dominant eigenvalues and eigenvectors of the system that, in mathematical terms, can serve to illuminate both analysis and action. Identifying them, using our empirical data sets, will be fundamental to making sense of these co-evolving systems. It is thus that we will harness the mathematics of co-evolving systems to the empirical study of social dynamics and demonstrate its value to social and policy studies.

## 6. DESIDERATA FOR EMPIRICAL CASE STUDIES

Against this conceptual background, we now lay out the *desiderata* for the datasets we will require, when we study non-linear dynamics within co-evolving systems, using the Jain and Krishna model. We will then be able to judge whether the datasets available to us, for any empirical case study, sufficiently meet these *desiderata*.

### *COMPONENTS*

- A directed network of nodes, supporting populations with interdependent levels of fitness
- A 'fast dynamic' that plays out the interactions among nodes and reveals the levels of fitness they each enjoy
- A 'slow dynamic' which eliminates the least fit nodes and introduces novel replacements

[10] Thus for example, in nuclear reactors and their control systems, engineers monitor a dashboard displaying (in effect) the eigenvalues and eigenvectors of the system, as they shift over time.

*THE DATASET*

- Operationalise nodes, edges and 'fitness' by reference to the database and over a succession of time periods
- Specify how the fast and slow dynamics will be observed within the data
- Recognise the database as a social artefact, the 'arts of civilisation' applied to the social world, and specify how the database will be exploited in this regard.

*THE EMPIRICAL ANALYSIS*

***For each time period:***

- Construct the adjacency matrices of the network (depending on the database, this could be at various levels of granularity).
- Examine the network for the presence of autocatalytic sets (ACS).
- Measure the levels of fitness of the different nodes and examine whether these are higher for nodes involved in an ACS.

***For successive time periods:***

- Map the eigenvalues and eigenvectors of the network and how they change over time
- Evaluate the eigenvalues and eigenvectors as revealing different types of innovation, and as 'weak signals' of impending change

*THE SCHOLARLY AND POLICY DEBATES*

- Consider the significance of these findings for research debates on innovation and change, within the fields in question
- Assess the utility of such analytical methods for social scientists, and their practical value for policy actors, as 'weak signals' of impending change and as tools for monitoring interventions
- Assess the implications for other case studies we will undertake, researching dynamic co-evolving systems across other empirical areas.

We now turn to our case study of **patents**, as revealing processes and patterns of technological innovation.

## 7. PATENTS AS TECHNOLOGICAL INNOVATIONS

***Introduction***

Technological innovation is a key feature of the modern world, driving all else, as celebrated by such diverse luminaries as Adam Smith, Marx and Schumpeter. How might we study the dynamics of technological innovation, as innovators combine existing and new technologies in novel ways? What sort of datasets might allow us to track these changes over time and better anticipate their direction?

Databases of patents have developed over the last century; in recent times, they have been harmonised internationally. Many scholars have used them for studying different aspects of technological innovation; there is a rich and self-critical literature on which we can now build (Napolitano et al, 2018).

The World Intellectual Property Organisation (WIPO) is the forum charged with overall governance of the international patent system; it includes the major national patent offices and the European Patent Office (EPO). WIPO is responsible for the global patent classification system and its annual updates; EPO is responsible for publishing the global PATSTAT database, set within the latest version of the classification system.

A patent can be viewed from two standpoints (Strumsky et al, 2012). First, it constitutes a new capability, a *force of production*. It combines and applies knowledge and technologies in new ways, whether incremental or more radical. This combinatorial ontology is central to much of the innovation literature, including Schumpeter and Hayek, and more recently Teece and Potts. [11]

Second, a patent is a claim to novelty; a novelty whose commercial potential the inventor wishes to protect. A patent application acknowledges the 'prior art' on which it builds, its intellectual debts, but it also makes clear what is new. A patent thus constitutes a claim to intellectual property from which others will be excluded - part therefore of the *relations of production*. Once a patent is granted, the inventor may develop it commercially. It may alternatively be sold, with the new owners either exploiting or shelving it, for a period at least, to avoid it threatening their existing product lines and markets.

As both forces and relations of production, patents remind us that technological change takes place within particular legal and institutional settings**.** These vary between countries and over time; they are socially and politically constructed and contested (Polanyi, 1944). Relevant here is the rich literature on 'varieties of capitalism' and 'national innovation systems' (Hall and Soskice, 2001; Lundvall, 2006). A study of patents should illuminate the co-evolution of technologies and institutions in different political economies. This does not mean that these are mutually insulated national domains. On the contrary, the patent regimes of individual countries can have consequences for innovation elsewhere, not least through the interconnections woven by international companies and their production chains.

It follows that a study of patents should illuminate major theoretical and empirical questions that are central to sociological enquiry. It should lay bare alternative trajectories of development and the scope for intervention by policy makers and corporate actors, in pursuit of those alternatives. We hope thereby to demonstrate that models of co-evolving dynamics deserve greater attention from the social science community at large.

### *Components of the Network*

We are applying to the social world a model of connected elements that co-evolve in a Darwinian system. Jain and Krishna have provided a mathematical model and explored its dynamic properties; we seek to apply this to technological change, using the patent datasets at our disposal. We hope to produce novel insights into the process of technological change. We hope also to illuminate the way in which big players – corporations and government in particular – use their power and position to shape the terrain on which innovation unfolds.

We have obtained privileged access to the PATSTAT datasets, which incorporate details of patents registered each year across all major patent regimes globally. We make selective use of these datasets to establish the components of our network, as follows: [12]

***Nodes:*** The PATSTAT database allocates each patent, newly registered in a given year, to a particular class and sub-class of its overall classification system – a dendrogram that is readily searchable. These **classes and sub-classes** we take as the **nodes** of our network. We can then study our network at different levels of granularity, ranging from the overall classes down to finer distinctions among sub-sub-classes etc.

***Edges:*** Each patent that is registered records other patents on which it draws and to which it is thus indebted: the 'prior art' by reference to which its own distinctive novelty can be viewed. The PATSTAT

[11] This contrasts with much of orthodox economics, which assumes a production function at the frontier of technology, shifting in response to technical progress, by reference to which businesses assess the profitability of different production mixes, as technology takers rather than technology makers.

[12] This particular application of the Jain and Krishna model to the PATSTAT data was originally proposed by Zeppini (2017).

database thereby links each patent (and the class and sub-classes within which it sits) back to the classes of those other patents, from whose prior art the patent in question benefits. This is a flow of know-how from those earlier patents and their patent classes to the latest innovations. We take these **citations** as the basis for building the **edges or links** of our network, representing flows of knowledge among patent classes.

***Observing the Fast Dynamic***

In the Jain and Krishna model, the processes of interaction among nodes that unfold during the fast dynamic drive the populations that thrive at each node and thus their **fitness.** In our case study we take this as given by the number of **new patents** registered in a given class (node) during a given time period. [13]

We expect those nodes having the greatest fitness to be associated with the overall direction of innovation and with the most innovative zones of the technology landscape. This is central to the investigation described below.

One problem might however appear rather challenging. In the Jain and Krishna model, the fast dynamic allows the interactions among the nodes of the network to proceed until they have reached an equilibrium, where we can see the final population of each node. Only then are we in a position to apply the algorithms of the slow dynamic, extinguishing one of the least fit nodes and adding a new node, a random mutation. Only after that, can we then run the fast dynamic again, on this modified network.

It is not clear how we would recognise that the patent network has reached such an equilibrium. Nevertheless, this is not as serious a difficulty as might at first appear, once we can with some confidence identify the emerging dominant ACS, along with its eigenvalue and eigenvector. This will allow us to see on what equilibrium the system is converging; it will also reveal the least thriving technologies and patent classes and thus the nodes that are candidates for extinction.

More generally, identifying the dominant ACS will reveal what we might call the ***zone of intensive innovation*** and the ***zone of stagnation***. This will (see below) illuminate the overall dynamics and direction of innovation.

***Observing the Slow Dynamic***

In the Jain and Krishna model, the algorithms of the slow dynamic extinguish one of the least fit nodes and add a new node, a random mutation. Only then do they run the fast dynamic again, on this modified network.

In the real world of patents, new nodes (in the form of new patent classes and sub-classes) appear more frequently. Every year WIPO publishes an updated version of the patent classification system, with EPO publishing the latest PATSTAT database, using that updated classification system. Both will typically include more than a few novel classes and sub-classes.

If the introduction of new nodes is here much more plentiful than the frugal mutations allowed by Jain and Krishna, the elimination of patent classes is in comparison much more grudging. Few patent classes or major sub-classes are ever wholly eliminated from the classification system, they are just left to stagnate. It is, however, right that our research should focus primarily on the new zones that open up and the zones from which they primarily draw knowledge. Patent classes long sterile can

[13] An alternative would have been to take the *total* number of patents in a given class (not just the newly registered) as our definition of fitness (Zeppini, 2017, section 1.3); but this would distract attention from the zones of most intense innovation with which we are most concerned: see section 8 below.

reasonably be ignored, whether or not the WIPO classifications and the PATSTAT database retain them. [14]

For Jain and Krishna, new edges are associated with the newest nodes. It may be objected that in the real world of technological innovation, new edges are often associated with existing nodes. We might explore some corresponding mathematical variant of the Jain and Krishna model. Nevertheless, the sense of the model is surely this: we should focus on the new nodes and edges that drive change more generally, and without which the existing system will tend to stasis.

Again therefore, it seems reasonable to keep calm and carry on. Our main focus is quite properly on zones of Intensive Innovation - zones where the annual addition of new nodes (classes and sub-classes) and edges is most lively. We can investigate how far these are home to the dominant ACSs – and whether, by watching the eigenvalues of the system, we gain insight into the emerging zones of change.

***Grasping a Contested Social Construct***

Now however we face a rather different a sort of challenge. We see and discuss objects using the concepts and language available to us, not least in regards to new technologies. Such concepts are socially constructed and contested. This does not imply that those technologies have no reality of their own - a reality which (although mediated by those concepts) we may attempt to explore and analyse. Nevertheless, we must take account of the social processes by which these concepts are constructed and modified. This is all part of what we earlier referred to as the 'arts of civilisation' – involving human reflection, experimentation and the growth of knowledge, all set within a struggle for positional advantage.

The international patent system is one of those processes. An inventor registers a patent, to establish a claim of intellectual property, but it may then take several years, before the patent is granted. During this period the inventor is in negotiation with the patent office, seeking an agreement as to what is new, about this particular patent - how the novelty should be conceptualised and classified, within the overall patent system. That system is itself however in flux.

WIPO publishes its annual revision to the classification system, adding sub-classes here, or merging sub-classes there, or even adding whole classes elsewhere. This ensures that the classification system keeps up to date with major reconfigurations of technologies; it allows inventers to search more easily for the prior art, when they are making new applications for fresh patents.

The EPO also publishes annually its latest PATSTAT database of patents worldwide every six months. This publication re-allocates patents across the revised classification system, so that they can all be viewed by reference to this most up-to-date picture of the technology system, as it exists today. This includes not only the new patents that inventors applied for over the past year, but also the patents from previous years that were already on the database. An older patent may now be re-assigned to a class or sub-class different from that which it originally occupied; and to one, indeed, that may have been only recently created. [15]

---

[14] A technology (class) such as the steam engine may at some point cease to be a zone of innovation and no longer attract any new patent applications. This is not because the steam engine has necessarily exhausted its potential for innovation, but rather because other technologies and the dynamics of economic development have diverted innovation and investment to other areas – this in response to new openings that have appeared, in part out of the successes of the steam age. It should never however be assumed that those old technologies will never have new applications – see for example windmills and ceramics, re-invented for the modern age. Thus technological development can open up new technological vistas which cannot be wholly predicted – and which may indeed appear rather like a random new node arriving.

[15] Within the DCICSS project, Napolitano and Pugliesi (2017) have created a unified dataset (PATSTAT 1400) to map each patent to the codes it had under different classifications in different years, so as to track not only classification changes but also the detailed flow of patents in this ever-changing classification. This uses raw

This however presents us with significant conceptual and methodological challenges, as we seek to map our network of nodes and edges and the dynamics in which they have been involved over successive years. This is especially the case for areas of technology that are zones of innovation – the most interesting for us – because it is here that the patent officers will in general have been most active in modifying and updating the system of classification.

## 8. PATENTS AS A CONTINGENT HISTORICAL PROCESS

This conceptual and methodological challenge opens up an important opportunity for our analysis of innovation. Far from the selection of patents as our case study being ill-advised, it will prove to have been particularly apt.

We have thus far been interested in how technologies co-evolve with each other, in autocatalytic processes as modelled by Jain and Krishna. Nevertheless, we grasp technologies through the institutions by which we organise them. The pace and direction of innovation depends, indeed, as much on the new institutions that are being invented – laws on e-commerce for example and on new forms of intellectual property – as on the new technologies themselves. The designers and inventors of these institutions (in government, in business, etc) are re-shaping the world, no less than the designers and inventors of technologies.

In this way, **institutions** and **technologies** interact and co-evolve with each other. Many new technologies emerge, only to find themselves on an institutional terrain that stifles them. In some other country, the institutional terrain may be more supportive. Some technologies and institutional forms may thus discover an 'elective affinity' that enables both of them to thrive and together to dominate their socio-economic ecosystem. [16] These *elective affinities between technologies and institutions* are therefore as potentially significant, as those between different technologies, in generating autocatalytic effects and in driving processes of innovation.

The international patent system is one such institutional domain; its officers are one group of institutional inventors. How then shall we think of the autocatalytic dynamics involved in the development of new technologies and their registration as patents?

Consider **Figure 7.** We start from the **north-east quadrant** and proceed clockwise.

A new year has started. WIPO has updated the classification system it published 12 months earlier, involving a dendrogram of classes and sub-classes. EPO has published the latest version of the PATSTAT database, locating all patents, both recent and not so recent, within this most recent version of the classification system. It thus maps the technological capabilities of the society. It also records the knowledge claims that are officially recognised, and the flows of know-how on which successive generations of patents relied - all organised in as parsimonious, searchable and up-to-date a form as possible. This mapping will now provide the stable framework that inventors can use, to search the 'prior art' and to register their new patents over the coming months. It will also allow patent officers to evaluate those new applications.

> *We ask: What technological capabilities do these patent classes capture - and what flows of know-how, from established patent classes (and sub-classes) to new patents, did they involve?*

---

data provided by the EPO's PATSTAT officers. It involves relevant data extracted from versions of PATSTAT published between 2007 and 2017.

[16] This co-evolution is not entirely blind. The inventors of new technologies pay attention to changes under way in laws on IPR, e-commerce etc. This will influence their decisions as to where they invest their time and creativity; and as new institutional spaces are created, they may shift their focus. Meanwhile, institutional inventers watch what new technologies are emerging, when they consider how to modify the legal and administrative environment.

*We construct an empirical network, with **patent classes as nodes** and **knowledge flows as edges**.*

We return to this NE quadrant, after working our way through the other three.

This moment of publication is the starting gun for the registration of new patents. Within **the <u>south-east quadrant</u>**, inventors start registering new patents and populating the classes of the newly-updated classification system. They are using the latter to record the prior art and register what is novel about their invention. They cite a variety of existing patents (and thus patent classes) as the antecedents on which they have drawn in novel ways.

This is the *continuous present of the technological fast dynamic*. Within this quadrant, the classes are given and fixed, just as with the nodes in Jain and Krishna. Inventors anchor their claims to novelty within this stable substrate, the prior art; it is only thanks to the parsimonious clarity and searchability of the classification system, that the novelty of their own invention can be rigorously demonstrated.

This is *the technological fast dynamic –*

*We ask: How do the various nodes of the network (in this case, the classes and sub-classes) thrive differentially? What is their resulting **fitness** – the number of patents registered in each class or sub-class during the year in question?*

**Figure 7 Patents as CHM**

| | **INSTITUTIONS** | | **TECHNOLOGIES** |
|---|---|---|---|
| ***The slow dynamic – a single moment every 12 months*** | ***Institutional slow dynamic*** **Patent office simplifies the system of classification - so that it is parsimonious in terms of classes and cross-citations across old as well as more recent patents. It reassigns all patents to their appropriate place within this amended classification system** | → | ***Technology slow dynamic*** **The amended system of classes is published, for inventors to populate with new patents** |
| ***The fast dynamic – a continuous process over 12 months*** | ↑ | | ↓ |
| | ***Institutional fast dynamic*** **Pressure mounts on patent offices to make ad hoc adjustments to the classification system** | ← | ***Technology fast dynamic*** **(New) patents populate (new) nodes** |

Even as the inventors register their grounding in the prior art, they also document the novelty of what they offer. These novelties will in due course, and taken together, undermine the existing order. Here therefore is a stable classification system, but one across which there progressively spreads a mass of prospective instabilities. In this way, the continuous present unfolds on the stable ground of the past, but also defies and challenges it.

In places, the sheer numbers of new patent applications threaten to overwhelm the categories' capacity to describe and contain the promised novelties, without blurring or losing their distinctiveness. In other places, the novelties defy the categorical boundaries and the distinctions they draw. This makes it increasingly difficult for the national patent officers to apply the classification system as it stands. We might indeed describe the pressures for such revisions, and their distribution across the array of patent classes and sub-classes, as the rate of activity – the fertility - of different parts of the *classification system*. This is the *institutional fast dynamic*, located in the **south-west quadrant.**

This is a highly creative and expert activity, a practical craft on the part of the national patent officers, as they consider what adjustments would be needed to provide all these new patent applications with an appropriate niche, capturing and displaying its distinctive novelty. As in all scientific work, such classification is an essential tool in building conceptual links for understanding and progressing the architecture of knowledge. It goes beyond mere categorisation and search; it allows each patent to come to the notice of other relevant strands of invention, enabling cross-pollination and the discovery of further elective affinities. Nevertheless, these micro-processes are largely invisible; we cannot view them directly within the patent publications; we can do no more than notice their aggregate net effects at year-end, by observing the changes that have been made in the annual updates of the classification system. [17]

Now, in the **north-west quadrant,** we move *from the fast to the slow dynamic*; from the video of the continuous present, to looking back from the end of each year, a careful snapshot, a *pause for constructive reflection*.

WIPO reads the reports and recommendations from the national officers. It takes stock of the pressures they have encountered for adjustment of the classification system and the conceptual tangles these have produced. It asks how to simplify these and to display the tangled web of citations as economically as possible. In information theory, this parsimony appears as a counterpart to the export of entropy in physics. Elements of the classification system serving no useful purpose are eliminated, as surely as the least fit species in an ecosystem.

We might distinguish between new patents registered in a given year, using the classification system published at the start of that year, and patents registered in earlier periods, but now re-classified by reference to the current system. The *ad hoc* modifications made during each year, by the local patent officers within the south-west quadrant, are driven primarily by the former. The latter also matter however; and when the international office considers what annual revisions to make, it will need to weigh up how easily those earlier patents can be mapped into the revised classificatory matrix that the newest patents suggest, while still maintaining a parsimonious and searchable scheme. [18]

---

[17] Lafond has been one of the scholars who has made some empirical study of these processes: Lafond and Kim (2017), Lafond (2014). He notices that they follow a strict process of checking and search, when they evaluate any new patent application, and search for the relevant prior art. He notices also that faced by a flood of new novelties, they tend to adopt a pragmatic two-fold approach: on the one hand adding extra sub-sub-classes, to permit finer distinctions; on the other merging (sub-)classes into a completely new class or sub-class, when sufficient novelties put the distinction between them in question. The first of these involves an 'incremental' change to the classification system, while the second involves changes that are to some extent more 'radical'. Given our interest in understanding the co-evolution of technologies and institutions, a fuller and more rounded enquiry might include study of the empirics of this *institutional fast dynamic*.

[18] Revisions to the classification scheme may be postponed, if they risk rendering it more complicated and less searchable. (In the same way, the addition of extra epicycles to the Ptolemaic picture of the universe seemed necessary at the time, to deal with new astronomical observations, but only produced a more complicated

This is the institutional slow dynamic - with WIPO looking back and checking the consistency and elegance of the various ad hoc adjustments that their local counterparts have proposed on the basis of the previous twelve months. The new classification of technologies that they now publish - and the new version of the PATSTAT database that the EPO then publishes – is the result of this reflection. The slow dynamic is not just a look back at the preceding year, simplifying and pruning the tangles of the classification system, after the myriad local adjustments it has suffered over the past months. It also looks forward and establishes the new classificatory terrain for the year ahead. This is the **north-east quadrant,** the slow dynamic in relation to technological forms. The new classification system will provide the stable framework that inventors will use, to register their new patents and their debts to the prior art, over the coming months.

As researchers, we can read the newly published WIPO classification scheme and the PATSTAT database and we can compare them with their predecessors of 12 months earlier. We can construct a modified empirical network, involving an amended array of classes and sub-classes as our nodes, and of the knowledge flows captured in citations, as our edges. By comparing it with its predecessor of the previous year, we can see what new knowledge debts characterised this latest cohort of patents.

We can then proceed to identify the emergence of autocatalytic sets and the eigenvalues and eigenvectors of the technological fast dynamic, as viewed from the standpoint of the present classificatory system.

## 9. A WELL-CHOSEN CASE STUDY

The previous section displayed the co-evolution of technologies and institutions, patents and classifications, in a contingent historical process. Centre-stage are the zones where a proliferation of new technologies is manifest in patents that radically challenge the established classification system. It is the dynamics of this zone of intense innovation, in both technologies and institutions, that we seek to capture, using our models of co-evolving networks taken from Jain and Krishna.

Consider now the relationship between **Figure 7** and **Figure 3**.

**Figure 3** offered a parsimonious representation of co-evolutionary dynamics. It was the starting point for our presentation of the Contingent Historical Model – a radical alternative to the General Linear Model, which has dominated much of social science. It looked back from the ecosystems of today to the elective affinities that had developed, among the populations of different species, during earlier periods. These affinities enabled some species to thrive, while others failed and became extinct. Only by revisiting those earlier periods, was it possible to make sense of the complex food webs that we see today.

**Figure 7** has taken us on a similar journey through successive time periods, in order to make sense of the patent system we see today, with a wide range of interrelated technologies, classified to clear and concise effect. The journey alternated between the slow and the fast dynamics, and between institutional and technological innovations, as we took transverse cuts through successive annual data sets. Those cuts involve alternating snapshots and videos. The fast dynamic tells of a continuing

picture of the heavens: Kuhn (1970)). At the very least, WIPO will want to take its time. They seem to have found that their 12-month cycle is about the right periodicity for the technology innovation system: allowing the classification to retain its freshness and relevance, while also not rushing to adopt every modification that local patent officers might suggest.

present; the slow dynamic looks back from the end of each year, a pause for both reflection and foresight. [19]

We thus disassemble the problem into four moments, each of which we study by appropriate interrogation of the PATSTAT databases for successive years. With the data laid bare, from this succession of vantage points, we can analyse the eigenvectors and eigenvalues of the system and illuminate thereby the autocatalytics of change. The selection of patents as our case study thus proves to have been particularly apt.

Jain and Krishna are concerned with the elective affinities, the positive synergies that develop among species, the mutualistic co-evolution. That is also what Figure 3 displayed. We must not however forget **Figure 4**. That was concerned with antagonistic co-evolution – the disruption that one species may cause for another and the pressures for the latter to adapt or be extinguished. Here the directed edges among nodes are negative not positive – and the mathematics of the Jain and Krishna model are rather different.

## 10. CONCLUSION

Using the PATSTAT data, Napolitano et al (2018) have published preliminary findings on the autocatalytic origins of innovation. Their study incorporates many elements of the conceptual framework described here. They demonstrate that the evolution of the technology network involves a growing autocatalytic structure. They also show that in the core of the autocatalytic set, the technology fields display greater fitness, in terms of a greater number of patents. Finally, they observe core shifts, whereby different groups of technology fields come to dominate the autocatalytic structure, only then to be overthrown. This points to radical innovation, with new combinations developing among distant as well as closely related technology fields. These are just the sorts of dynamic change that Jain and Krishna explore through analysis of eigenvectors and eigenvalues.

This demonstrates the promising possibilities for this form of analysis. This however will need to be set within a wider institutional analysis, including for example the literature on different national innovation systems (Lundvall et al, 2016) and varieties of capitalism (Hall and Soskice, 2001). It will also need to integrate co-evolutionary dynamics with an analysis of power, political economy and the struggle for positional advantage (Room, 2015).

This research also has potential benefit for policy makers, aiding their strategic imagination and providing indicators - 'weak signals' - of impending change, so that they can better steer the dynamics of autocatalytic innovation (Room, 2005).

We have used the Jain and Krishna model of co-evolutionary dynamics as the basis for our empirical social research. In the present paper, the focus was on new technologies, as captured in patent databases. That work is still ongoing and incomplete. Beyond the present programme of work, the Jain and Krishna model could, in principle, be applied to other fields of social, economic and technological change. In those future case studies we may need to modify the model, depending on their specific features; and there are always likely to be trade-offs between the realism of the model and its mathematical tractability. It is also likely that in many potential and promising areas of study, there are no appropriate databases. Nevertheless, this paper may at least have established the conceptual and methodological coherence of such an enterprise and its potential empirical application.

[19] This is close to the historiographical question raised by writers such as Croce and Collingwood (Collingwood, 1942): how should we study a historical episode in terms of both its significance for actors at the time, but also in terms of its significance for us today?

## REFERENCES


Abbott, A. (2001), *Time Matters: On Theory and Method,* Chicago: University of Chicago Press

Blaug, M., G.M. Hodgson, O. Lewis, and S. Steinmo (2011), 'Introduction to the Special Issue on the Evolution of Institutions', *Journal of Institutional Economics*, 7(3): 299-315.

Bronowski, J. (1981), *The Ascent of Man,* London: Futura

Brown, G. and A. Langer (2011), 'Riding the Ever-Rolling Stream: Time and the Ontology of Violent Conflict', *World Development*, 39(2): 188-198.

Checkland, P. and J. Scholes (1990), *Soft Systems Methodology in Action,* Chichester: John Wiley

Collingwood, R. G. (1942), *The Idea of History,* Oxford: Oxford University Press

Coyle, R. (1996), *Systems Dynamics Modelling: A Practical Approach,* London: Chapman and Hall

Crouch, C (2005), *Capitalist Diversity and Change*, Oxford: Oxford University Press

Darwin, C. (1859), *The Origin of Species,* London: Wordsworth (reprinted 1998)

Dawkins, R. (1976), *The Selfish Gene*, Oxford: Oxford University Press

Flannery, T. (1994), *The Future Eaters,* New York: Grove Press

Freedman, A (2013), *Strategy: A History*, Oxford: OUP

Gilbert, N. and K.G. Troitzsch (2005), *Simulation for the Social Scientist (2nd Edition),* Maidenhead: Open University Press

Goodman, M and J. Thornston (2018), *Client Earth*: Victoria, Australia: Scribe

Gould, S.J. (1991), *Wonderful Life,* Harmondsworth: Penguin

Hall, P. A. and D. Soskice, (ed) (2001), *Varieties of Capitalism: the Institutional Foundations of Comparative Advantage*. Oxford: Oxford University Press.

Hodgson, G.M. (2001), *How Economics Forgot History,* London: Routledge

Howe, R.H. (1978), 'Max Weber's Elective Affinities: Sociology Within the Bounds of Pure Reason', *American Journal of Sociology*, 84(2): 366-385.

Jain, S. and S. Krishna (2003), 'Graph Theory and the Evolution of Autocatalytic Networks ' in S. Bornholdt and H. G. Schuster (ed), *Handbook of Graphs and Networks* Weinheim: Wiley-VCH**:** 355-395.

Kaldor, N. (1972), 'The Irrelevance of Equilibrium Economics', *The Economic Journal*, 82(328): 1237-1255.

Kauffman, S.A. (1993), *The Origins of Order: Self-Organisation and Selection in Evolution,* Oxford: Oxford University Press

Kingdon, J.W. (1984), *Agendas, Alternatives and Public Policies* Boston: Little, Brown and Company

Koenig, M.D., S. Battiston, and F. Schweitzer (2008), 'Modeling Evolving Innovation Networks' in A. Pyka, A. Scharnhorst (eds), *Innovation Networks - New Approaches in Modeling and Analyzing* (Eds.) Heidelberg: Springer

Kristensen, P. H. and J. Zeitlin (2005), *Local Players in Global Games: The Strategic Constitution of a Multinational Corporation,* Oxford: Oxford University Press

Kuhn, T. S. (1970), *The Structure of Scientific Revolutions,* Chicago: University of Chicago Press

Lafond, F (2014), 'The Size of Patent Categories: USPTO 1976-2006', UNU-MERIT Working Paper Series: Maastricht.

Lafond, Francois and Kim, Daniel, 'Long-Run Dynamics of the U.S. Patent Classification System' (February 27, 2017). Available at SSRN: https://ssrn.com/abstract=2924387 or http://dx.doi.org/10.2139/ssrn.2924387

Lieberson, S. (1987), *Making it Count: The Improvement of Social Research and Theory,* Berkeley: University of California Press

Lieberson, S. and F.B. Lynn (2002), 'Barking up the Wrong Branch: Scientific Alternatives to the Current Model of Sociological Science', *Annual Review of Sociology*, 28: 1-19.

Lundvall, B.-A., P. Intarakumnerd, and J. Vang, (ed) (2006), *Asia's Innovation Systems in Transition*. Cheltenham: Edward Elgar.

Mayer-Schonberger, V. and K. Cukier (2013), *Big Data,* London: John Murray

Maynard Smith, J. and E. Szathmary (2000), *The Origins of Life: From the Birth of Life to the Origins of Language,* Oxford: Oxford University Press

Napolitana L and E Pugliese (2017), 'PATSTAT1400 PATSTAT Multi-version Combined Database' Institute for Complex System NRC

Napolitano, L., Evangelou, E., Pugliese, E., Zeppini, P. & Room, G. (2018), 'Technology networks: the autocatalytic origins of innovation' *Royal Society Open Science.* 5, 6, RSOS-172445.

Pierson, P. (2004), *Politics in Time,* Princeton: Princeton University Press

Polanyi, K. (1944), *The Great Transformation,* New York: Rinehart

Potts, J. (2000), *The New Evolutionary Microeconomics: Complexity, Competence and Adaptive Behaviour,* Cheltenham: Edward Elgar

Powell, J.H. and J.P. Bradford (1998), 'The Security-Strategy Interface: Using Qualitative Process Models to relate the Security Function to Business Dynamics', *Security Journal*, 10: 151-160.

Powell, J.H. and J.P. Bradford (2000), 'Targeting Intelligence Gathering in a Dynamic Competitive Environment', *International Journal of Information Management*, 20: 181-195.

Room, G., (ed) (1995), *Beyond the Threshold: The Measurement and Analysis of Social Exclusion*. Bristol: The Policy Press.

Room, G. (2005), *The European Challenge: Innovation, Policy Learning and Social Cohesion in the New Knowledge Economy*, Bristol: The Policy Press

Room, G. (2011), *Complexity, Institutions and Public Policy: Agile Decision-Making in a Turbulent World,* Cheltenham: Edward Elgar

Room, G. (2012), 'Evolution and the Arts of Civilisation', *Policy and Politics*, 40(4): 453-471.

Room, G. (2016), *Agile Actors on Complex Terrains*, London: Routledge

Schelling, T. C. (1978), *Micromotives and Macrobehaviour,* London: W.W.Norton

Sloan Wilson, D. (2008), 'Multilevel Selection Theory and Major Evolutionary Transitions', *Current Directions in Psychological Science*, 17(1): 6-9.

Squazzoni (2012), *Agent-Based Computational Sociology*, Chichester: John Wiley

Stewart, I. (1997), *Does God Play Dice? (2nd Edition),* Harmondsworth: Penguin

Strumsky, D, J Lobo and S van der Leeuw (2012), 'Using Patent Technology Codes', *Economics of Innovation and New Technology,* 21:3, 267-286

Toner, P. (1999), *Main Currents in Cumulative Causation: The Dynamics of Growth and Development* London: St Martin's Press

Zeppini, P (2017), 'Autocatalytic Networks of Technologies', University of Bath