The emergence of self-organisation in social systems: the case of the geographic industrial clusters

Andriani, Pierpaolo

The emergence of self-organisation in social systems: the case of the geographic industrial clusters

Pierpaolo Andriani

Thesis submitted in fulfilment of the Degree of Doctor Philosophy

University of Durham,
Durham Business School

2003

Abstract

The objective of this work is to use complexity theory to propose a new interpretation of industrial clusters. Industrial clusters constitute a specific type of econosphere, whose driving principles are self-organisation, economies of diversity and a configuration that optimises the exploration of diversity starting from the configuration of connectivity of the system.

This work shows the centrality of diversity by linking complexity theory (intended as "*a method for understanding diversity"*[1]) to different concepts such as power law distributions, self-organisation, autocatalytic cycles and connectivity.

I propose a method to distinguish self-organising from non self-organising agglomerations, based on the correlation between self-organising dynamics and power law network theories. Self-organised criticality, rank-size rule and scale-free networks theories become three aspects indicating a common underlying pattern, i.e. the *edge of chaos* dynamic.

I propose a general model of development of industrial clusters, based on the mutual interaction between social and economic autocatalytic cycle. Starting from Kauffman's idea[2] on the autocatalytic properties of diversity, I illustrate how the loops of the economies of diversity are based on the expansion of systemic diversity (product of diversity and connectivity). My thesis provides a way to measure systemic diversity. In particular I introduce the distinction between modular innovation at the agent level and architectural innovation at the network level and show that the cluster constitutes an appropriate organisational form to manage the tension and dynamics of simultaneous modular and architectural innovation.

The thesis is structured around two propositions:

1. Self-organising systems are closer to a power law than hierarchical systems or aggregates (collection of parts). For industrial agglomerations (SLLs), the closeness to a power law is related to the degree of self-organisation present in the agglomeration, and emerges in the agglomeration's structural and/or behavioural properties subject to self-organising dynamic.
2. Self-organising systems maximise the product of diversity times connectivity at a rate higher than hierarchical systems.

---

[1](Castells 2000) p.74

[2] Kauffman investigations

Acknowledgement

Every work of a considerable duration goes necessarily through ups and downs and periods of intellectual stagnation and doubts. A supervisor that keeps you on track is essential. If you can find a supervisor who can do by the simple art of showing confidence and trust in the supervisee, you have found the best. Some gentle nagging occasionally helps as well. Dick, thanks, I am indebted for that.

My sincere thanks to Alan Jessop for accepting the ungrateful task of stepping in at a late stage and providing useful advice and comments.

I am indebted to my friends at the Future Centre in Venice, especially Annaflavia Bianchi and Roberto Dallamura. Guys, thanks a lot. Some of my best ideas were conceived in discussion with you strolling along the canals in Venice. Without your support, generosity and friendship it would have been so much more difficult to finish this work.

Thanks to Mirella for her patient discussing my statistical doubts and to Joel for helping out with Mathlab programming.

Index

List of figures

List of tables

I hereby declare that the PhD thesis entitled: "The emergence of self-organisation in social system" the case of the geographic industrial clusters" has been written by me and has not been submitted for the award of any degree or other similar title previously

PREFACE

Complexity theory is a new interpretative framework that has the potential of leaving a mark on the evolution of thinking in natural and social sciences. It tries to explain the origin of order, the robustness and the decay (or catastrophic demise) of social and natural systems thanks to a few universal laws, which are more similar to dynamic trajectories of development than to the prescriptive laws of Newtonian sciences. Complexity is also a multilevel theory. Natural and social phenomena are the result of the aggregation of agents at multiple levels. At each level of aggregations new laws and new fields of endeavour emerge.

This thesis analyses industrial clusters using complexity theory concepts, together with other relevant theories and models. I consider industrial cluster as a specific type of complex system, based on distributed management of simultaneous innovation and production activities.

I start by telling *in anger* the story of the formation of one such system. Meadow, an isolated village based on "*backyard capitalism*" which evolves into a sophisticated cluster characterised by a complex self-regulatory production and innovation system, is the story of the passage of a system through a series of phase transitions (or bifurcations) each of which gives rise to the emergence of a different system structured around different dynamics. In short, it is the story of the emergence of order in a social system.

The example of Meadow is interesting as it posits the question: "If systems based on self-organisation emerge via semi-spontaneous creation of order in an endogenous way, then what theory do we need to make sense of emergence?" The phenomenon of emergence is difficult to analyse via theories that assume that the behaviour of the system can be reduced to a sum of its components' behaviour. The theory we need is one that a) takes account of the inherent limitation that emergence imposes on the knowledgeability of systems and b) embeds the emergence of new, more complex organisational levels into a more general theory that includes reductionism as a particular case.

This thesis formulates a theory that describes the emergence of industrial clusters around some general dynamic patterns based on mutually self-reinforcing growth and selection mechanisms. These mechanisms form a closed architecture of processes, which gives rise in an autocatalytic fashion to a specific dynamic identity of the industrial cluster.

The initial point of the theory concerns the relationship between diversity and growth "... *diversity probably begets diversity; hence diversity may help beget growth*"[3]. Diversity, writes Kauffman, is autocatalytic, it leads to further diversity. The expansion of diversity is likely to result in the closure of a catalytic cycle. In short, closure introduces internal rules of organisation, which causes the transition from an aggregate of parts into a system and the emergence of a coevolutive dynamic. Coevolution and closure are conjugated variables, which act to partially decouple the system from its environment. The system remains open to external information and energy, which are used to construct complex internal structures, whose organisational principles are however internally determined. In other terms the organisational rules are genotypic, whereas the space of implementation of the rules is phenotypic. There is currently no theory to describe this type of self-generative diversity. I have called it systemic diversity and defined it as the product of diversity times connectivity.

How can we recognise the effects of the expansion of systemic diversity around sets of autocatalytic loops? This question leads into the search for the structural characteristics that can be used as indicators of the specific dynamics described above. The starting point is that industrial clusters enjoy a set of properties: they consist of networks of autonomous agents; their pattern of connectivity exhibit weak and strong linkages. These features are connected. In fact, in self-organising systems chaotic and ordered features coexist. Chaotic dynamic allows for frequent reconfiguration and emergence of novelties, ordered dynamic allows for robustness and homeostasis. The balance between the two is described by the metaphorical expression of *edge of chaos*. At the *edge of chaos* the distribution of connectivity alternates strong links within highly connected sub-networks with weak links connecting the various local networks (within the system). The distribution of links within and across sub-networks and the dimensions of the highly connected sub-networks obey a power law. The emergence of a power law is a sign of a self-organising dynamic. In fact the power law indicates that the system tunes itself toward a state whereas the distribution of structural and behavioural

[3] (Kauffman 1995) p.292

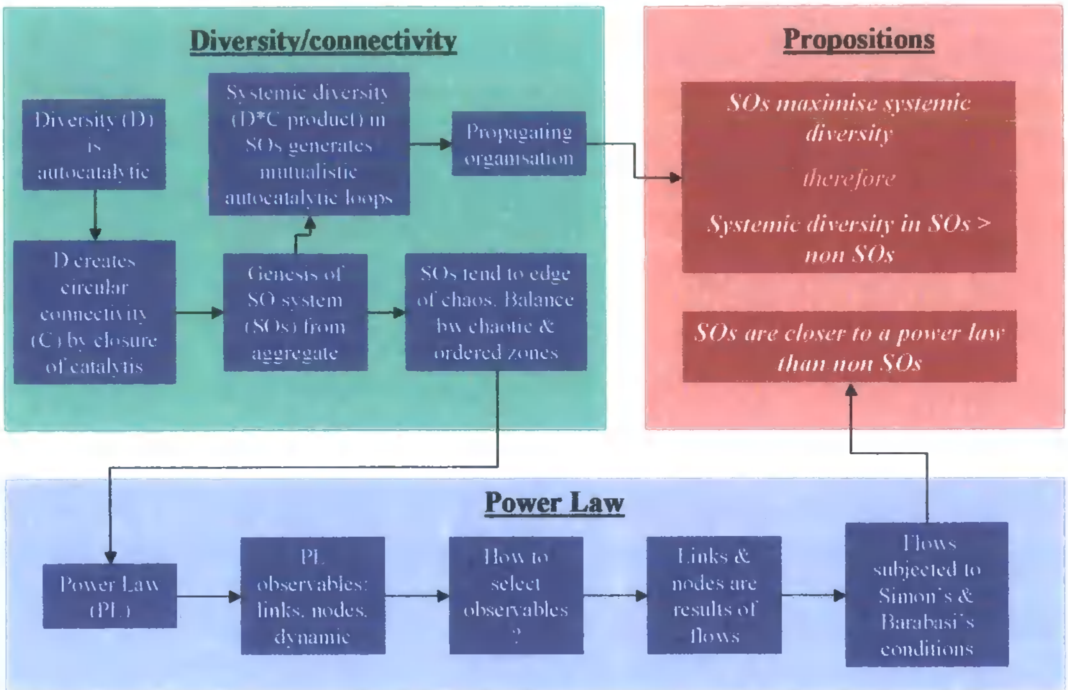
properties is such that all scales and behaviours of relevant network's properties are present in the system and follow a simple mathematical distribution. Power laws emerge in the three constituent elements of networks, i.e. links, nodes and behaviour. I have presented the argument, whereby the occurrence of power law for these three elements is the manifestation of the same underlying dynamic, that is, the self-organising nature of networks at the *edge of chaos*. I have also suggested that a transformation of variables changes one variable into the other, thereby raising the possibility of using any variable for the determination of the structural and/or dynamic state of the system. In short, self-organisation is explainable in terms of the set of interdependencies that determine the set of flows between nodes and links. This is something that a statistical analysis on appropriate systems can test. If self-organisation and power law are correlated then we can use power laws to reveal self-organisation in action. From this point of view, a comparison between systems characterised by different degree of self-organisation should reveal a correlation between self-organisation and closeness to power law.

The theory summarised above is used to generate two propositions:

1. Self-organising systems are closer to a power law than hierarchical systems or aggregates (collection of parts). For industrial agglomerations (SLLs), the closeness to a power law is related to the degree of self-organisation present in the agglomeration, and emerges in the agglomeration's structural and/or behavioural properties subject to self-organising dynamic.

2. Self-organising systems maximise the product of diversity times connectivity at a rate higher than hierarchical systems

The logical steps regarding the translation of the theory into propositions are represented below.



The following step concerns the validation of the theory summarised above. In the chapter on methodology I first, discuss the methods of research adopted to test the main propositions, provide a rationale for those methods and illustrate the limits of the research methods; second, I introduce the units of analysis and provide a rationale for their selection; and third, I expose the procedures and steps taken to validate the propositions. This chapter offers two main contributions. First, I propose and implement a method to measure the effect of self-organisation in industrial agglomerations by using a power law distribution. Second, I propose and develop a method to measure generic and systemic diversity in industrial agglomeration.

The empirical work, carried out on Census data relative to different types industrial agglomerations in Italy, confirms the validity of the propositions. The picture that emerges is as follows:

1. A power law analysis is an effective instrument to discriminate between systems driven by self-organising dynamics and systems driven by hetero-organised dynamics (aggregates of parts and hierarchical systems). In the context of the Italian industrial agglomerations my results confirm that the self-organising dynamics that characterise the presence of a cluster can be revealed by a power law analysis. Therefore my suggestion that a power law is an indicator of the emergence of a system from an

aggregate of parts is confirmed by the empirical data. Moreover, as industrial agglomeration are a complex mix of self and hetero-organised dynamics, I show that a power law analysis is able, by making use of a regression analysis, (and, at least in a statistical sense) to indicate the weight of the different dynamics.

2. The point made above together with first, the idea regarding the equivalence of the different power law theories and second, the discussion of the correlation between self-organised criticality and edge of chaos dynamic, suggests that a power law distribution is an indicator of an *edge of chaos* dynamic. Industrial clusters then are social systems on the *edge of chaos*. When such a dynamic arises, the system shows a power law distribution of chaotic and ordered features. As this type of distribution doesn't have a typical scale, this means that industrial clusters thrive and survive by exploring a very broad range of possible dynamics.

3. The previous point leads me to the second part of the results. The origin of the self-organising dynamic, revealed by the power law analysis, lies with the capability of exploring diversity. My results confirm that self-organising systems maximise a particular type of diversity, that I called systemic diversity. I propose to measure systemic diversity by the product of diversity (as measured by a mix of three parameters, variety, balance and disparity) and internal connectivity. I also propose to use the concept of systemic diversity as a way to operationalise Kauffman's concept of the *propagating organisation*[4].

4. Specifically, my research demonstrates that industrial agglomerations closer to the cluster form show a higher systemic diversity than systems more dominated by aggregations or hierarchical organisational forms. Moreover, in terms of the characteristics of the Italian industrial agglomerations this study has further refined the taxonomy proposed by Cannari and Signorini[5] and generally confirms the validity of their approach. However, the large distance that separates industrial agglomerations of type D1 (so called super-clusters) from all other types of industrial agglomerations suggests that a discontinuity separates the two groups. This result seems also in line with the autocatalytic nature of cluster formation and development,

---

[4] see next section

[5] (Cannari and Signorini 2000)

which would exclude a continuous distribution of agglomeration types between the cluster and non-cluster.

What are the final results of this research? Clusters constitute a specific type of econosphere, whose driving principles are self-organisation, economies of diversity and a configuration that optimises the exploration of diversity starting from the configuration of connectivity of the system. The maximisation of diversity and connectivity takes place "*without destroying the accumulated propagating organisation that is the basis and nexus from which further novelty is discovered and incorporated into the propagating organisation*"[6]. The structure of the propagating organisation obeys a power law distribution. This distribution maximises the exploration of nodal features, connectivity patterns and system wide behaviours.

These results confirm the general intuition by Kauffman that the evolution of biosphere and econosphere is subjected to some general 'laws' that cut through the specifics of individual systems. Power laws and the idea that organisations are propagating set of autocatalytic processes expanding diversity along trajectories set by environmental pressures and internal connectivity are two of these 'laws'. My research develops a specific formulation for these two laws in the economic context of industrial agglomerations and demonstrates their validity.

---

[6] (Kauffman 2000) p.85

Chapter 1 - Introduction - Increasing returns and local economies: the story of Meadow

I will tell an imaginary story in order to illustrate the concepts of: economies of diversity, autocatalytic sets and dynamics of co-evolution. The story concerns a village called Meadow which, over a long period of time, develops milk and other food related products. The story will show the power of co-evolution to speed up the process of knowledge creation and innovation when coupled to the mechanisms of economies of diversity. We will see that the concept of clusters and therefore the whole history of Meadow is counter-intuitive. Much of the classical economic tradition is based on the distinction between large and small firms. The asymmetry principle states that a large company can do everything that a small firm can do, but not vice versa. This principle and the literature of economies of scale, learning curve effects, monopolistic control of market, effectiveness in controlling intellectual property rights issues and in influencing standards setting, all of these have skewed the industrial policy agenda toward the principle that big is automatically good. However clusters of small firms, which act as a system, stand as an alternative model of organisation able to compete under conditions of high market uncertainty. The cluster creates value exploiting self-organisation of internal configuration and the self-sustaining loops of economies of diversity. I have condensed in the following story some typical mechanisms of cluster formation and evolution as they have been encountered in the Italian districts and to some extent in the American literature on clusters, highlighting the mechanisms by which diversity acts as a catalyst of innovation and innovation pushes forward the boundary of existing diversity, thereby generating a self-sustaining loop. The story is used to illustrate the dynamics of formation and survival of industrial clusters (that I define synthetically as economies of diversity, see section 3.2.2) and to generate most of the research propositions that will be tested in the rest of the thesis.

1.1 THE BEGINNING

Let us imagine a village, called Meadow, in which a simplified economy operates around farming and breeding, a type of economy that Krugman would call "*backyard capitalism*"[7]. Many households own farms, others offer labour. Most farmers own a few cows and oxen from which they produce milk and meat for their own immediate consumption (Figure 1).

One day somebody made the serendipitous discovery that heating milk very quickly helped to make it last longer.

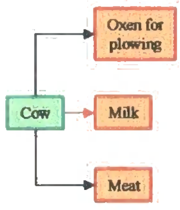


Figure 1 - The beginning

"*Chance favours the prepared mind*"[8], said a famous local man of knowledge, referring to the fact that this discovery had occurred several times in other places, but only in Meadow it was transformed from household into commercial use. At first the innovation was used as a provision for winter. This initial intuition was soon followed by other opportunities of putting the invention to good use. A nearby land called Pasture had been devastated by war. Cows had virtually disappeared. The first 'export' took place along family links (to provide support to distressed relatives or long time friends). A part of the population (the "*conservatives*") refused to use the mysterious milk and their leaders argued for the ban of import and consumption. It was wicked, they said, in an attempt to block one of the very early examples of globalisation. However UHT milk

---

[7] "*...Suppose that we really lived in the constant returns world that much economic theory still assumes. Then it would be hard to understand why the economy is not characterised by 'backyard capitalism', in which each household or small group produces most items for itself*" (Fujita, Venables et al. 1999) p.2)

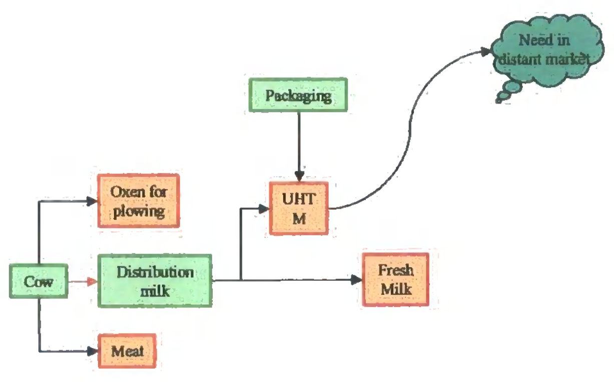
[8] Famous sentence by Luis Pasteur

spread along underground channels and contributed to the establishment of commercial (and ultimately cultural) links between Meadow and Pasture.

The export activities represented a big change for Meadow. The establishment of permanent trading links, first with Pasture and then with other nearby places, caused a quick increase in trading activities and the establishment of the first proper firm, Cow.

In addition to fresh milk Cow diversified into fresh and long lasting milk, the former for local consumption and the second mainly for export.

The process of transporting and distributing long life milk required appropriate packaging, which in turn required the development of complementary assets[9] (packaging, transport, distribution), without which the initial innovation could not have been commercialised. With the rise of export over time production and distribution tended to grow apart. This is an example of the synergistic effect of not purely substitutive innovation (Figure 2): development of a new product and diffusion of innovation caused a series of cascading effects to existing sectors (i.e. transformation of the productive structure of Cow with the spin-off of Distribution) and birth of synergistic new sectors (i.e. packaging)[10]



[9] See (Teece 1987) for a discussion of this point

[10] Synergy indicates that when two activities become complementary, their relation becomes regulated by some sort of feedback loop. In the specific example the feedback is positive, that is the more product is sold, the more packaging is needed and, the more product is sold, the larger the income to improve and innovate the packaging sector.

Figure 2[11] - Appearance of related activities

Let us sum up: an accidental discovery led to a flurry of economic, cultural and social changes in Meadow. This discovery opened a technological trajectory[12] of development[13] populated with unintended consequences[14].

Our imaginary story goes on with another discovery. One day whilst disposing of the interiors of a lamb, a piece of stomach fell into the pot where milk was being heated. The day after the interesting observation was made that the milk had turned into a relatively hard, gelatinous substance, with a sour-like but overall pleasant taste.

---

[11] In the figures, firms are shown in green and products in red. New products and new firms are highlighted by filling the boxes.

[12] The concept of technological trajectory was elaborated by Giovanni Dosi (Dosi 1988). Dosi's theories and models are based on the Kunhian concept of scientific paradigm and on evolutionary economics theories (Nelson and Winter 1982). Ultimately the concepts of paradigms and technological trajectories rest on the use (and translation) of biological evolutionary models to the history of technological change.

[13] Although the concept of technological trajectory seems to point towards the direction of determinism (a road connects A to B along a predetermined path), in reality the example given (Meadow) and the evidence coming from the history of technological change indicate that the relationship between the macro-invention (the one that opens the technological trajectory) and the multiplicity of micro-inventions (see (Mokyr 1990)) for a complete discussion of the terms) that follow are determined by the local conditions in place. Because those local conditions can never be pre-stated, as many of them are ephemeral, it follows that technological determinism, at least in a strong version, is not applicable.

[14] The idea that the history of innovation is largely an history of unintended consequences is presented in the soft edge ((Levinson 1997)). For instance: the beginning of packaging, the separation of Cow from Distribution, etc.

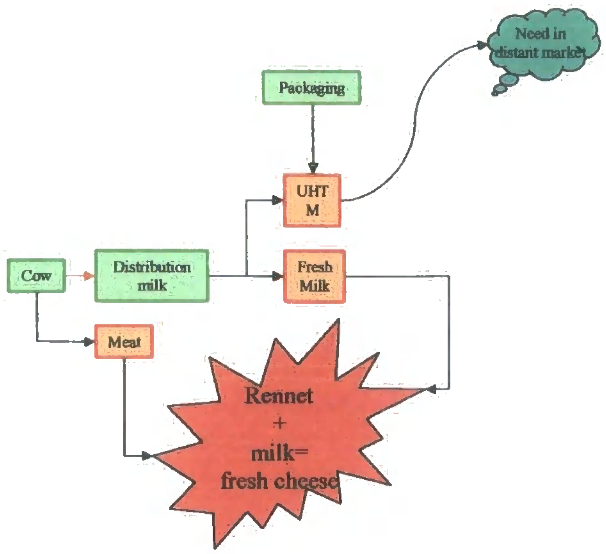


Figure 3 - Complementarity-based innovation

A set of unplanned circumstances favoured the discovery of cheese and opened a new market and industry. In particular the integration and spatial proximity of the activities that led to the discovery of cheese created the conditions for serendipity to play its role. Yet in order to convert this discovery into an innovation it was necessary to isolate the production parameters from the random circumstances in which the discovery had taken place and to identify a market.

Resolution of problems in Meadow had little to do with a traditional scientific approach. Instead it was very practical, new combinations of parameters would be probed and the successful ones would be selected. This '*probe and learn*'[15] approach would slowly build a stratified castle of largely experiential knowledge.

The initial production of cheese was at the beginning (before the isolation of rennet and the optimisation of production methods) slow, cumbersome and diffused. Many households undertook the production of cheese for internal consumption, some to add some variety to the existing diet, others for sheer curiosity. The result was a 'collective' effort (multiple parallel experimentation) that generated several local approaches (recipes) to cheese making. In fact, different initial sets of parameters defined different starting points for different technological approaches to cheese making. These were

---

[15] (Lynn, Morone et al. 1996)

made of many incremental improvements and based on the tacit knowledge acquired through successive waves of many mistakes and few successes[16]. When somebody got by trial and error a recipe that worked, that recipe was soon captured into a set of rules. However, the experiential and tacit nature of the technical process implied that the various recipes were partially incompatible with one another. This fact, together with the multiple experimentation, created a pool of different approaches that contributed to the diversity of the local system.

Unlike long life milk, the discovery of cheese was path-dependent, that is, it required the presence of some necessary circumstances to take place. The historical aspect of path-dependency set Meadow on a path of geographic differentiation[17] from adjacent territories. In particular, the presence of a specific technological trajectory based on milk and its derivatives, a favourable attitude to experimentation and innovation, a set of channels to commercialise products, and especially the optimism coming from previous successes made Meadow different from other territories and contributed to trigger a spiral of accumulation of knowledge, resources and wealth.

The discovery of cheese required a previous degree of variety in the environment, in this case the tangible and intangible factors such as milk and meat preparation, heat, the right mental framework to notice the discovery and to exploit its significance. The innovation took place at the convergence of these fundamental factors and exploited some unforeseen effects of that convergence. The consequence of the discovery was the creation of a link between some previously unrelated activities. This resulted in a net increase in a) the variety of the system, (more final products) and b) its complexity (because the new product was a by-product of existing products it generated a network of interdependencies among the system's variety).

The discovery of long life milk and cheese gave rise to the process of coevolution[18], whereby the feedback of interdependencies among economic activities determined the lock-in of the system around specific technological trajectories and the emergence of a collective system of value creation. In particular the two innovations generated a) the

[16] Optimisation of a single set of initial conditions is analogous to climbing a local peak. Experimentation with multiple sets results in a multi-peaked environment (See (Kauffman 1995))

[17] In the parlance of complexity theory a symmetry breaking

[18] This concept will be explored in more details in the following chapters

formation of a seed of an industrial and commercial network, b) the transition from a closed to an open system and c) the creation of relevant knowledge about how to manage innovation and its consequences. The agents constructed their own environment and in so doing they not only affected the material and cultural conditions according to which the processes of survival and change take place, but also contributed to define the space of future innovations[19].

The organisational transformations that followed the discovery of cheese are fascinating. The production of cheese during the initial experimental period was carried out by a fluid and integrated organisation. The increase in knowledge (about the cheese making processes) allowed the definition of the interfaces between sets of related processes, causing the emergence of distinct production phases based on specific, largely self-contained tasks, around which integrated production routines could be aggregated. The set of competencies, people and technical tools in each phase constituted a module. The emergence of modules caused also the shift from product to process innovation[20]. With time the organisation fragmented internally into groups to generate in a second phase specialised spin-offs.

Modularisation made possible the spin-off of the rennet production activity. This spin-off however did not create an island isolated from the rest of the industrial activity in Meadow, but, on the contrary, due to the strong complementarity between Rennet and Cow, increased the pattern of collaboration and flow of information between the two firms. The interdependence between their products forced the two companies to cooperate with each other and made them mutually interested in each other's well being (Figure 4). It also constituted the first example of unbundling, that is, the separation of newly acquired functions from the existing ones via the spin-off mechanism. The long lasting effects of this original experiment will be examined in the next paragraphs.

---

[19] The relationship between agents and environment is circular. Agents construct relevant features of their environment through their own micro-actions and the environment selects what actions are possible and provides a context for those actions.

[20] See (Abernathy and Utterback 1978) for a discussion of this point

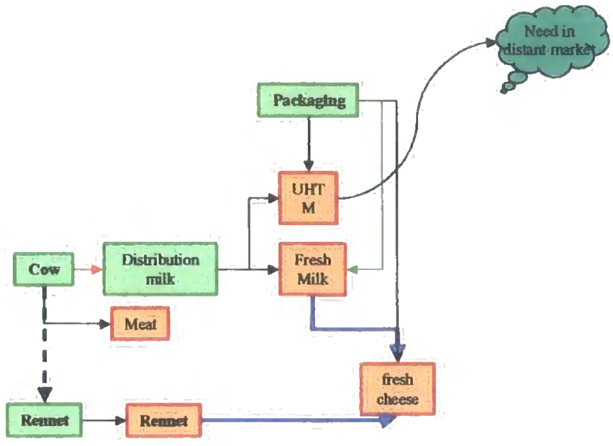


Figure 4 - Spin offs and partnerships

The initial diffusion of cheese was local, but the recent experience of long life milk and the interest that the news of the invention of cheese had raised in the distant land of Pasture convinced the group at Cow that a way could be found to make long-life cheese. It was suspected that the solution for long-life cheese could be similar to that of long-life milk. And indeed the thermal mechanism was successful. The new product was called parmesan.

This innovation presented few striking features:

1. Random circumstances played no role. For the first time the expansion of the product base was planned in response to a demand formulated by the external market. It was an early example of market-pull innovation

2. Knowledge creation was largely incremental and based on adaptation of an already existing and well working knowledge base[21].

The formation of interdependent chains of activities and the appearance of synergistic complementary activities, such as packaging, posed a problem of coordination as changes in one channel (such as the allocation of resources) affected the performance of the other channel (Figure 5). What was the correct balance between producing milk and meat? How did the balance ultimately affect the production of cheese and other derived

---

[21] (Hall and Andriani 1999) introduce the concept of substitutive and additive knowledge. Also relevant on this matter is the concept of absorptive capacity (Cohen and Levinthal 1990): this describes the likelihood of acceptance of a new piece of knowledge in function of the existing knowledge base and existing mechanisms for related knowledge creation and diffusion

products? The coordination problem could be solved either by internalising the activities within one centrally managed company or by delivering control to autonomous agents and letting the players find an acceptable level of exchanges in a trial-and-error way (and ultimately) self-organising way. The system faced a bifurcation[22] as both ways were potentially able to provide an acceptable solution. The spinning off of Rennet from Cow however pushed the system toward the autonomous agents solution.

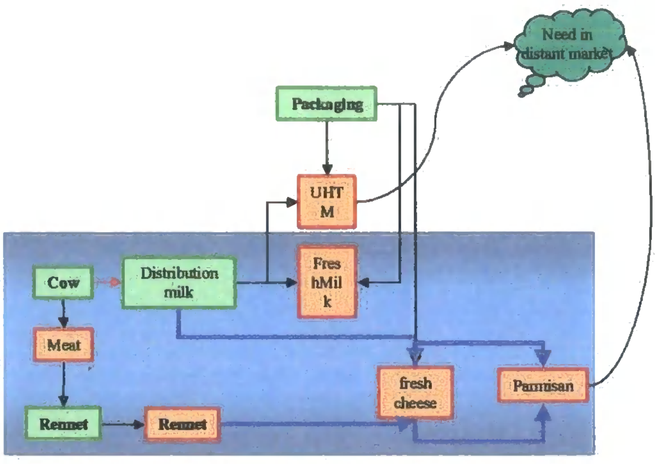


Figure 5 - Appearance of socio-economic loops: from chains to networks

The story goes on. The economic development of Meadow triggered a series of social changes. As often happens with family owned companies, the second generation was given the possibility of a more formal education. The fact that the first generation identified in education an asset for the betterment of their companies was a result of the fact that the history of Meadow had largely been based on the exploitation of new knowledge.

The first case of knowledge driven innovation was the discovery of casein. One of the youngsters observed that the final process of milk fermentation was a protein, which was subsequently called casein. The discovery was again serendipitous, but its understanding and use required scientific and not only experiential knowledge. Scientific knowledge is a form of theorising from experience that distils the common content between different phenomena and allows prediction of similar but not yet tried phenomena[23].

---

[22] (Allen 1997) p.18

[23] See Seely Brown on philosophy of knowledge (Brown 2000);(Popper 1974), (Kuhn 1996)

The discovery of casein in itself would have been confined to a laboratory curiosity had not somebody made the observation that cows actually liked the protein and further observed that casein seemed to be an effective source of energy and increase milk and meat production. This observation is not trivial as cows are vegetarian. It required a subtle mental shift[24]. In accepting that cows could also eat food of animal origin, our inhabitants learned that the categories around which the world seems to be organised were not fixed but could, to a certain extent, be manipulated in order to exert control over the natural environment and to modify it to one's advantage.

The discovery of casein led to an experiment that had long-lasting consequences. Cow's products presented a marked seasonal pattern. When the winter came in, and the grazing fields were covered by snow, production went down. Casein provided a partial solution because it could be stocked during periods of high production and used as food during winter (Figure 6).

The loop milk-casein-milk was markedly different from the cheese loop. In the latter the interdependency between the two chains posited a problem of co-ordination. In the former the production of a component led to a higher production of the following product, that in turn fed back on the original component in an autocatalytic[25] way, each component being at the same time input and output of the loop. The extraction of casein from the treatment of milk made possible the increase in production of milk from which casein is extracted making the cycle (to a degree) self-sustainable. Interestingly, although the single phases of production related locally to one another (previous and following step), the collective cycle was auto-referential. When taken in isolation, each step of the loop (dubbed as hypercycle, see note 18) added value to the following one. However the closure of all the steps around the cycle provided the organising principle to the

---

[24] Kuhn in the 10th chapter of the Structure of Scientific Revolution reports the following: "*The Chinese, whose cosmological beliefs did not preclude celestial change, had recorded the appearance of many new stars in the heavens at a much earlier date. Also, without the aid of a telescope, the Chinese had systematically recorded the appearance of sunspots centuries before these were seen by Galileo and his contemporaries. ....The very ease and rapidity of with which astronomers saw new things when looking at old objects with old instruments may make us wish to say that, after Copernicus, astronomers lived in a different world*" (Kuhn 1996) p.117

[25] Eigen and Schuster received the Nobel Prize for the discovery and theory of hypercycles, or autocatalytic sets

collective system. The autocatalysis[26] loop behaved as a relatively closed system, in which the role of the single parts was defined as a function of the collective property of the system, and at the same time the collective property emerged as a result of the interactions of the individual parts. Unlike the production of cheese, the organising principle behind the autocatalytic loop was intrinsic to the loop itself; it was the loop. The organising principle corresponded to the architecture of the loop and both (principle and architecture) were nothing else than the collective function carried out by the autocatalytic set.

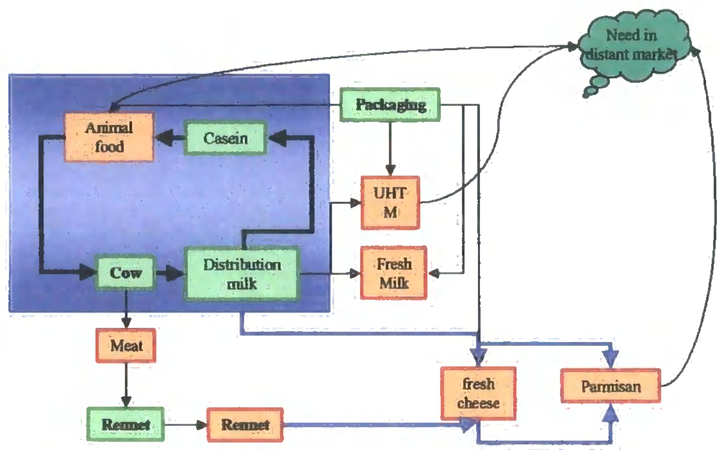


Figure 6 - Emergence of autocatalytic set

The discovery of casein and the application of casein to autocalytically increase the production of milk in winter had a strong effect on revenues. The new market of artificial animal food was created which became a further source of export.

Once started, the autocatalytic cycle of animal food production became stronger and exerted a strong centripetal attraction. A truly "increasing returns" dynamic was in action here. The more casein was available to feed the animals, the higher the production of milk and meat, which in turn made easier to produce more casein. The self-sustaining dynamics of the loop caused an unprecedented period of rapid industrial expansion in Meadow.

Another relevant aspect has to do with the type of industrial organisation that the autocatalytic loop generated. We saw in earlier examples that different organisational

[26] For a discussion of autocatalytic sets see (Kauffman 2000) and (Juarrero 1999)

structures were compatible with the production of similar products. The self-sustaining and centripetal logic of the autocatalytic loop instead favoured the centralisation of the activities of the loop within Cow. The necessity to defend the secret of the casein, the nature in which activities fed and sustained each other, the sequentiality of the cycle and the relative ease with which activities could be centrally coordinated all played a role in increasing the centralisation of activities within Cow. The closure of the set of activities provided a natural boundary for this part of the organisation. What had, until then, been a story of relative decentralisation, took a twist toward centralisation and the dominance of a single player.

Figure 7 shows the situation in Meadow. We see two main groupings of activities: the autocatalytic loop around casein and the cheese loop. Packaging was ancillary to both.

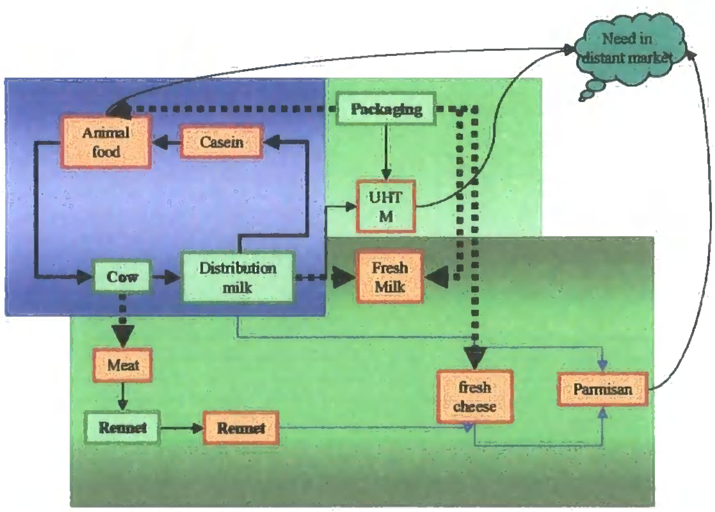


Figure 7 - Mutualism between catalytic sets

The historians of Meadow wondered what would have been the industrial structure of Meadow, had the autocatalytic cycle of casein been discovered in an earlier time, when the structure of production and society in Meadow were simpler. A look at **Figure 8** will help make the point. Let's imagine that the wheels of history could be turned back and that the discovery of casein had taken place before the discovery of cheese and parmesan.

Under this condition the rapid increase in wealth subsequent to the discovery of the autocatalytic set of the casein would have frozen the configuration of the industrial structure in Meadow. The increasing returns situation would have made a single player, Cow, reap all the benefit of a *de facto* monopoly. The swift increase in wealth and the resulting monopolistic position would also have created the illusion that Cow had achieved the right structure and business model that ensured a sustainable competitive advantage. Arguably the stimuli for further innovation would have been stifled. Provided that the stream of revenue kept coming, the best energy of Meadow would have been absorbed in the expansion of the autocatalytic cycle and penetration of further markets. The likely evolution of Meadow would have been determined by a strategy of specialisation and exploitation of the benefits resulting from the autocatalytic cycle.

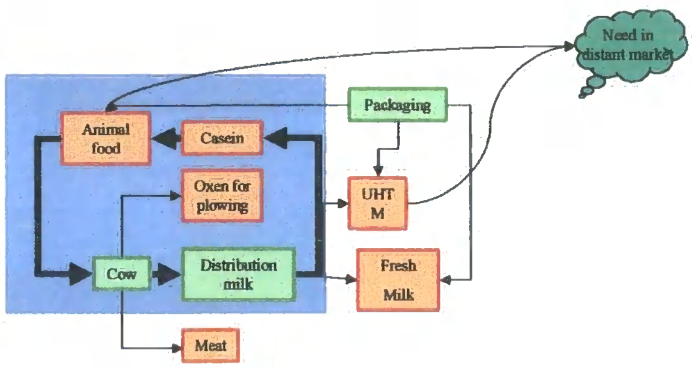


Figure 8 - Reversing the wheels of history: lack of diversity and dominance of autocatalytic set

Summing up: an early discovery of a mechanism that allows a fast accumulation of resources and wealth in the context of a relatively homogeneous local environment freezes the structure and the future of the organisations around that of the winner. In absence of environmental shocks, the system follows a trajectory of incremental improvement based on of the exploitation of existing resources around the dominant business model (specialisation strategy) in the pursue of efficiency. As this strategy is characterised by increasing returns, this causes the emergence of a monopolistic player. However, the specialisation around a single production model makes the system more vulnerable to environmental changes. The search for efficiency streamlines the organisation eliminating redundancies and slack, hence reducing the internal diversity of the system, thereby reducing its adaptability.

In any case the mutual relationships between loops and actors in Meadow prevented the collapse of the industrial structure around the scenario seen above. The diversity of the system acted to defend its internal differentiation. Again this was not the outcome of any conscious, planned or rational decision. Cow did not see any particular advantage in acquiring Rennet or disrupting the effective partnership with Packaging. Those lines of business expanded at a much slower rate than Cow's core business and were therefore left alone.

1.2 LOCATION ATTRACTIVENESS

Since the beginning of the industrialisation of Meadow around the business of milk and its derivatives, the village had grown bigger and richer and its social structure had become more complex and stratified. A powerful group of merchants and industrialists had emerged and several figures related to the new industry, commerce and export had developed. The energy for the development and expansion of industry and commerce in terms of labour force and entrepreneurial energy had substantially come from within Meadow. However, the success of Meadow generated interest in the communities nearby and in all the territories served by Meadow's products. It seemed to them that the inhabitants of Meadow had hit upon a magic recipe to become rich. With time Meadow became a symbolic place where fortunes could be created. This generated an interesting positive feedback loop[27]. Businesses and people moved to Meadow to establish their businesses or to find a better job. The consequent diversification of its economic and social base created further opportunities for business expansion and employment of workforce, thereby increasing Meadow attractiveness for business and talents inflow. The changes that this brought were radical. The best workforce started commuting to Meadow and would-be entrepreneurs started investing there. The first people to arrive were either relatives of locals, friends or people that had some relationship with Meadow. Their arrival was welcomed and their integration speedy. Workforce was much needed to support the expansion of the economy. After the initial arrival of people homogenous with the local culture, a second wave of more heterogeneous people coming from more distant places arrived. The rather surprising fact was the relative ease with which the integration proceeded[28]. The immigrants arrived in familiar groups, which held strong familial ties with each other. The integration proceeded in steps. At the

---

[27] See (Fujita, Venables et al. 1999) for a discussion of location attractiveness

[28] For a thorough discussion of a successful historical case of expansion by integration of immigrants in an industrial district, see Becattini ((Becattini 1997))

beginning the newly arrived started as workers. After learning the '*mysteries of the trade*' and after being exposed to the local entrepreneurial culture, when a favourable occasion arose, they tried the entrepreneurial adventure. They used social capital funding[29], that is, the saving put aside by the enlarged family. This mechanism provided several benefits:

- Access to financial resources within the local community;
- Financial screening of the entrepreneurial ideas provided by the members of the family, some of whom were presumably working in a similar industrial environment;
- Extreme commitments of the group to the success of the enterprise: if it proved successful, then a chance existed for the other members of the extended family to improve their material conditions of living.

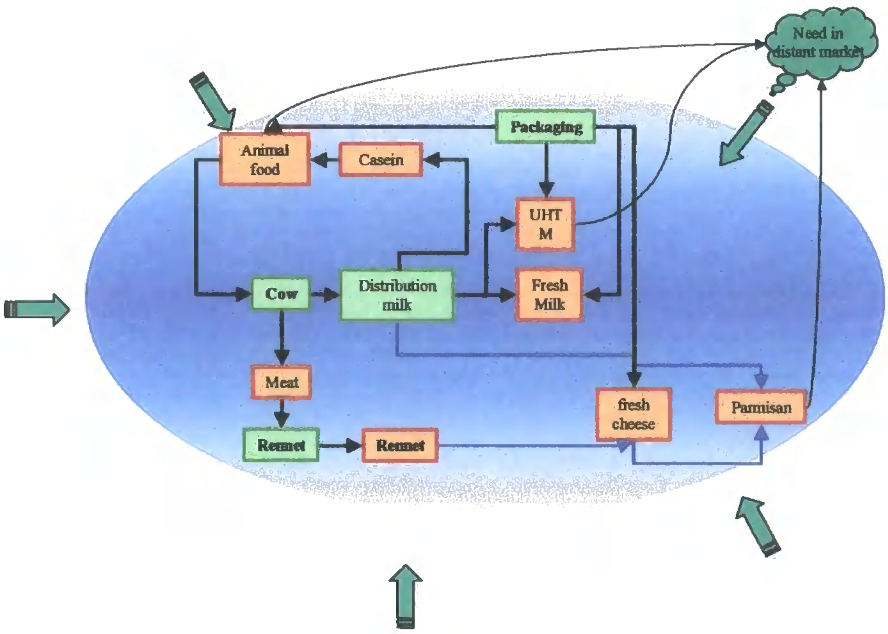


Figure 9 - Location attractiveness 1: immigration from nearby areas

At the end the capability of smooth integration of immigration within the social fabric of Meadow was due 1) to the favourable economic conditions that required the expansion of the local population and 2) to the pragmatic culture of tolerance, then common in Meadow. Instead of segregating the immigrants to menial jobs, the local

---

[29] for a discussion on social capital see (Coleman 1990; Burt 1992; Putnam 1993).The concept of the social capital bank as an alternative means of promoting entrepreneurship is again in (Becattini 1997)

culture valued the contribution that bearers of new cultures could provide and allowed these new groups to exploit the possibilities that emerged as a result of the fusion of different cultures. The final effect was a powerful increase in social diversity, which was nonetheless socially interacting with the extant social body[30].

There is a second aspect worth mentioning. The influx of newcomers brought knowledge and competencies previously unknown in Meadow. Some part of this knowledge had never found an industrial usage, because it had never before found a fertile ground. For example, lack of sophisticated customers (with a taste for innovative products/services), absence of complementary assets[31], lack of credit, etc. would prevent the transformation of knowledge into innovation. All these features, were present in Meadow in a concentration higher than anywhere else. Meadow was an *innovative milieux*[32], a place where new ideas had a higher probability of finding complements and linkages that could spark creativity and innovations. It is a bit like a Lego construction kit. Having the basic pieces in place allows a certain type of models to be realised, let's say a simple box. But, if somebody manages to make wheels, axles, and planks, then a simple carriage can be realised. Wheels, axles and planks are objects built with simple Lego pieces, but they perform a superior level function, which no single piece in isolation can perform. In this way we get a hierarchy of parts, which can achieve progressively more complex functions. First, single pieces, then, simple subsystems (the wheel) and then systems (the cart). The combinatorial space of possibility (let's call it the grammar[33]) increases exponentially at each level. A *milieux* is a place where the construction of levels of combinatorial complexity has achieved a more sophisticated

---

[30] It will be a central point of this thesis to indicate that diversity becomes autocatalytic if the diverse actors do interact with one another. In this case diversity favours the emergence of systemic properties.

[31] See (Teece 1987), complementary assets are tangible and intangible assets, such as mass manufacturing, miniaturisation capability, access to specific knowledge capability, superior logistics, etc., that may make the difference between two innovators with the same idea

[32] see (Castells 2000) for a definition of *milieux innovateur*

[33] The concept of a combinatorial grammar has been elaboratored by Walter Fontana a theoretical chemist at the SantaFe Institute (in (Kauffman 1995)). A similar concept having to do with the space of possibilities generated by higher order enabling constraints is in (Juarrero 1999)

level than in other places. This allows for serendipity[34] to happen. An example is shown below.

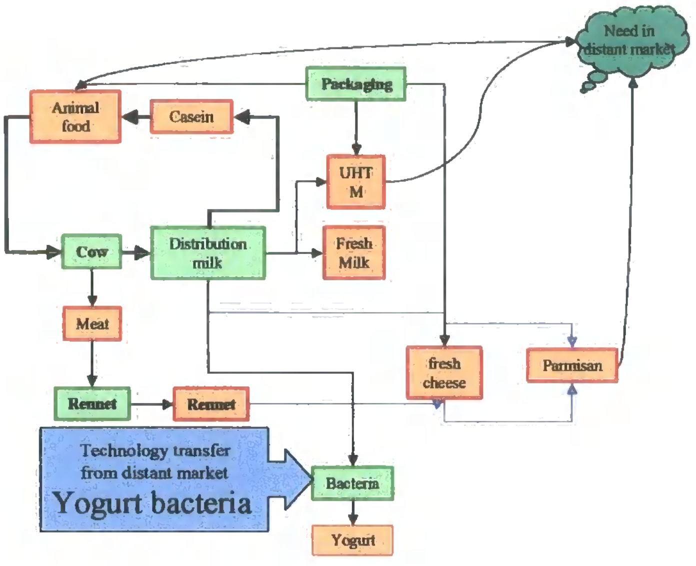


Figure 10 - Location attractiveness 2: technology transfer and increase in diversity

In some villages it was known that under certain circumstances the fermentation of milk would produce a creamy, lumpy substance that had the property of self-regenerating if given the appropriate nutrient (milk). For historical reasons this substance was used for its medicinal properties and its technique of production had never diffused outside. It was also known how to keep it going but not how to start it all over again. The immigrants from that area brought it with them. In a densely connected and highly diversified environment such as Meadow, information circulated quickly and the news of the wonder that this substance could do spread around. By transferring a pattern of usage into a new environment, it is very likely that a form of adaptation will take place. The people of Meadow, not constrained by the immigrants' dominant mindsets, tried new uses for the new substance. They started using the new substance as food with associated medicinal properties, thereby unconsciously starting the field of functional food.

At that point the usual mechanisms of collective experimentation started again. Through the use of scientific and experiential knowledge about milk and its derivatives,

---

[34] I have introduced in section 3.2.4 the concept of exaptation. This concept sheds light about the dynamic of serendipity

the problem of how to produce this substance was cracked and a production technique found. The name yoghurt was given to this new substance. The diffusion of yoghurt was relatively speedy as it was already in demand thanks to its medicinal properties.

This was probably the first case of technology transfer in Meadow. Although the discovery happened somewhere else (interestingly in a place less developed than Meadow), yet, yoghurt flourished in Meadow, where the right circumstances were in place[35]. One of the effects of location attractiveness was to act as an attractive vortex for 'dormant' knowledge. Awakening "dormant knowledge" required a fertile ground where first diverse technical and market competencies and second presence of innovators[36] could act as evolutionary selective forces promoting mutation, adaptation and selection of product forms. Without those forces yoghurt would have remained confined to a medicinal household use.

The consequences of the new discovery were multifold. In graph terms, when a new species is added to the population ecology, a new node is added to the extant network (**Figure 10**), with the consequence that new links are generated, whose number is roughly proportional to the total number of nodes (assuming complete connectivity amongst nodes). Because links are informational channels, the potential for information exchanges (and therefore innovation) increases quadratically[37] with the number of nodes. In short technology transfer acted in a twofold way. Directly, by stimulating innovation, and indirectly by increasing the structural depth and diversity of the existing network, thereby creating the conditions for further innovation (as it will be seen in the case of architectural innovation).

---

[35] *"what looks like chance may be as much the result of preexisting local circumstances"* [Porter, 1991 #412] p.239

[36] The field of diffusion of innovation defines innovators as the first segment of customers interested in novelties (see Rogers, diffusion of innovation (Roger 1995))

[37] This is known as Metcalfe law (Shapiro 1999)

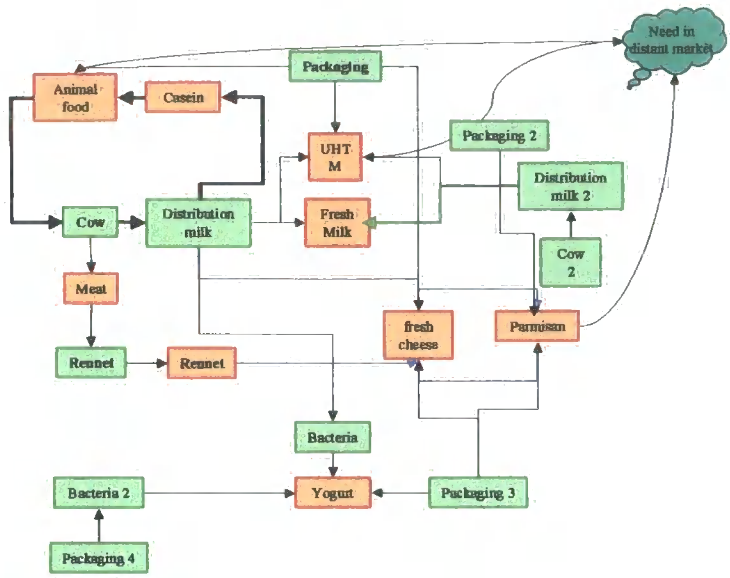


Figure 11 - Location attractiveness 3: internal competition

The story of Meadow has been so far based on diversification and exploration of emergent synergy. However, imitation is a powerful evolutionary mechanism that works to generate internal competition. During phases of economic expansion there is room for "me-too" products. Actors without radical new ideas but with a sense for business opportunities exploited existing products, either by duplicating them, or by improving the fine tuning between consumers demand and current products/services, using process innovation, branding, complementary services, etc. In this way they triggered the formation of additional firms in niches or sectors already covered by incumbents.

Another source of internal competition comes from the natural tendency of protected knowledge and trade secrets to diffuse outside the firms' boundary. Inventions get reinvented[38], magic recipes unveiled, complicated patterns of logistic and production operations replicated. So the best techniques of production of rennet, yoghurt, casein, etc. slowly stopped being the protected secret of few and became common knowledge within Meadow. "*The mysteries of the trade are as if they were in the air*"[39]. But they didn't diffuse outside Meadow. Single elements of the know-how did, but the complex set of technical, experiential and productive knowledge that had been accumulated over

---

[38] Reinvention is facilitated by the certainty that a solution of a certain problem exists. See Mokyr (Mokyr 1990) and the example of the reinvention of bone china

[39] Marshall, cited in (Maskell 2001)

time remained circumscribed in Meadow. The continuous iteration of steps, the progressive modularisation of production, the set of microscopic improvements made the knowledge available only to the people that possessed the vocabulary and the grammar to understand that knowledge. And because that vocabulary was tacit, it was very difficult to replicate. We can consider Meadow as the historical and structural context in which the single elements of knowledge acquired meaning.

For all the reasons mentioned above the rate of internal competition increased (see **Figure 11**).

Now there are in general two ways of competing: to do things better or to do better things. The result of the internal competition was two fold: it generated a pressure first, to reduce production costs and keep healthy margins (neoclassical competition); and, second, to diversify products in order to escape competition by generating new temporary monopolistic control in market niches (Shumpeterian competition[40]). **Figure 12** shows what happened in Meadow. The accumulation over time of several modules of related activities and products within a bundle of technological and product trajectories created ideal conditions for diversification of products based on the combination of existing modules. For instance, it did not take long to imagine that a sweeter yoghurt could appeal to customers that found the taste of yoghurt rather sour and that a potential endless variety could be generated mixing sugar or pieces of fruit to create a product at the interface between fruit and yoghurt. Other simple but very profitable innovations could be generated by mixing existing cheeses with certain fungi that were well known to confer a particular taste to food to invent blue cheese. Once the gates were opened to architectural innovation[41], the only limit to experimentation were a) the availability of modules to be picked and mixed, b) a ready-to-experiment local market, and c) availability of funding. All were present in Meadow. Architectural innovation became a

---

[40] See (Saviotti 1996). The distinction between rivalry and competition is a further reason behind the shift to Shumpeterian competition. As (Porac and Rosa 1996) write: "defining rivals is not so much a matter of overt behaviour as it is one of managerial attention and discrimination". Co-localisation provides the necessary cognitive dimension (Boari, Odorici et al. 2001) for rivalry to emerge from competition

[41] See (Henderson and Clark 1990). The use we make of architectural innovation is slightly different from H&C. They point out that architectural innovation is often at the beginning of a new technological trajectory. Many macro-inventions are therefore based on architectural innovations [Mokyr, 1990 #367]

combinatorial game, whereby the higher the number of interconnected modules, the higher the chance of finding complement-based new products.

The story of Meadow suggests a general law that may link diversity and innovation. Above a certain threshold, the more internally diverse the system is and the more rapid and dense the exchange of information within the system are, the higher the ratio of Schumpeterian to Neoclassical competition will be[42].

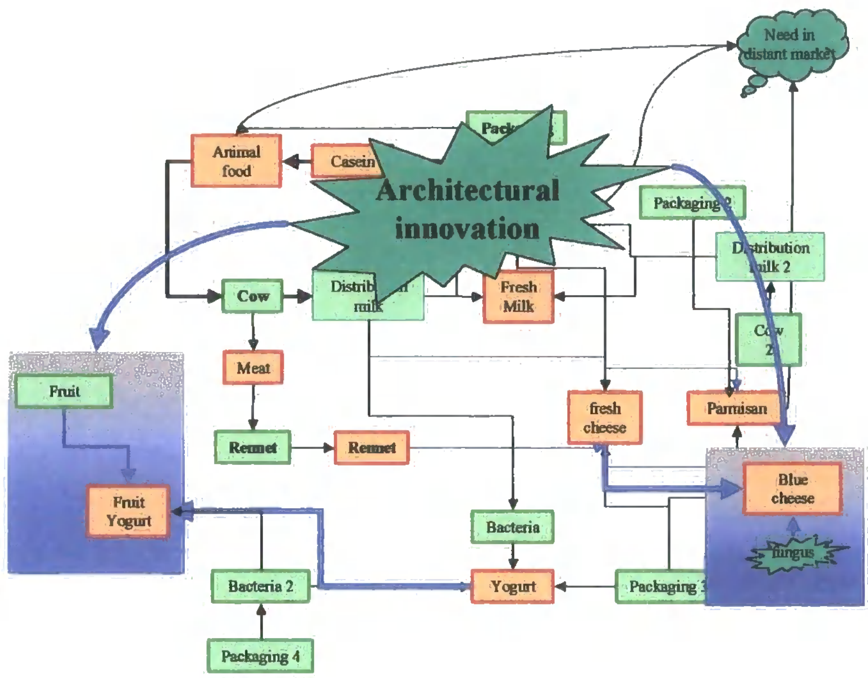


Figure 12 - Location attractiveness 4: Shumpeterian competition

1.3 UNCERTAINTY, DISINTEGRATION AND PHASE SPECIALISATION

The long phase of economic expansion caused an increase in the quality and number of activities within the main company. Cow, the first business that started with production and commercialisation of milk, had now spread its product range to cheeses

[42] The formulation of this "law" owns much to a) Bill McKelvey's idea that speed of coevolution favours niche selection of population selection effects ; b) Saviotti's distinction between Shumpeterian and classical competition (Saviotti 1996) and c) Kauffman's (Kauffman 2000) theory of the role of diversity in catalytic web of reactions in the origin of life

(in partnership with Rennet), casein and animal food. At the same time operations and support activities (legal, administrative, veterinary, export, cleaning, R&D, training, etc.) had grown increasingly more complex. Finally, with increasing specialisation, simple operations had become fragmented into a series of routines and procedures requiring admin and specialised support (for instance tool-making workshops).

Cow had really become a conglomerate with serious problem of management and co-ordination. The centralised structure that still dominated the company was a legacy of the early days. This has progressively caused the bureaucratisation of its management structure.

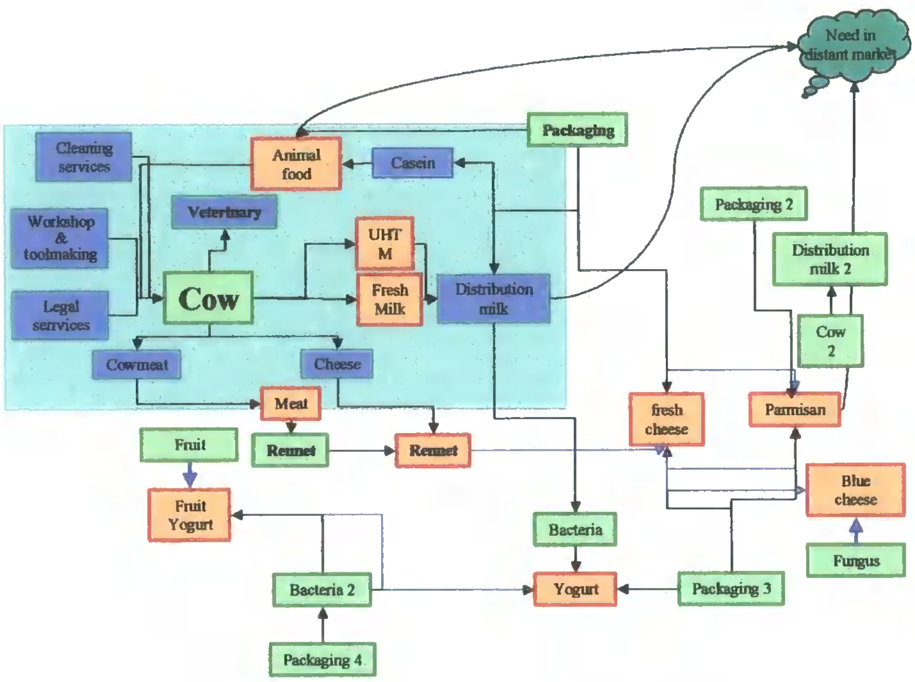


Figure 13 - Uncertainty, disintegration and phase specialisation

The structure began to disintegrate when a crisis hit Meadow. Some of the traditional markets of export had shrunk because of local circumstances, just when an investment euphoria had left the Meadow's households indebted. Although the contraction was relatively small compared to the sales volume[43], its effects were amplified by the consistent debt exposure. The situation generated even more uncertainty when some

[43] The term *butterfly effect* has been coined to describe a high sensitivity to initial conditions. See (Gleick 1987)

debtors could not repay their debts on time, triggering a domino series of failures. Owner managers tried to reduce their exposure to risks, via the usual means, layoffs, delays in debt repayment, costs and price reduction, reduction in inventory, etc. At the same time they tried to focus on the core business and get rid of ancillary and service activities, whose contribution to cash flow was hidden or negligible.

At the social level, the sudden increase in unemployment caused much social tension and the diffusion of communitarian ideas between the most affected social classes. But, the reliance on the social capital provided by the extended family structure and the multiple ties between social and economic activities supplied a safety net that ensured that the material conditions of living could be maintained at least for a while. Whether this set of multiple ties did stop emigration from depleting the accumulated intellectual and social capital in Meadow or whether it was a form of innate optimism that any crisis could be overcome, it is difficult to assess. At the same time some workers, especially the most intelligent and specialised ones, saw the opportunity to start their own activity. Instead of facing a brutal layoff, they proposed to their employers to lease some machinery related to non-core activities to be redeemed against future work. The worker would take control for instance of the lathe and repay the employer ensuring cheap supply of parts or maintenance. The contract made sense to the employer who would outsource a non core activity, keep control of the worker via the threat of calling off the informal deal, save money on future work (as the overheads on the micro-enterprise were lower) and finally maintain a relationship based on trust and previous history of successful collaboration with competent workers[44]. On the worker side the risk of starting an activity during a recession, supported by the extended family safety net, was balanced by the following factors.

1) The informal contracts between on the one hand the ex-worker and the employer and on the other hand the worker and its extended family allowed the start up to proceed without little need of funding from financial investors.

2) The business started with a portfolio of orders guaranteed by the well-known needs of the parent company.

---

[44] This mechanism is beautifully described by (Becattini 1997) in the history of Prato

3) Flexibility: during periods of low activity the worker could find temporary activity, for example, in the farming part of the extended family; during periods of high demand, his relatives could offer support and cheap labour.

4) Finally, the substitution of the relationship of employment with one of economic dependency (until the repayment of the machinery had been done) increased dramatically the potential range of actions of the worker allowing the micro-enterprise to extend its range of customers and products[45]. When the economic situation turned positive again, the micro-enterprises would start diversifying and looking for new customers and new ways to put to profit its competencies, thereby weakening the economic dependency but keeping the collaborative relationship.

All the causes mentioned above caused the disintegration of the conglomerate activities in Meadow and the real start of its clusterised industrial structure. In **Figure 13** it is shown how Cow disintegrated into a network of firms dealing either with services, such as cleaning, veterinary, distribution, etc., or with specialised aspects of the value chain, such as, cheese, casein, meat and tool-making. A similar process took place in other firms as well. The process acquired a self-sustaining dynamic. The sudden increase in the number of firms created further competitive pressure for product and market differentiation. The market dynamics changed as a consequence of the new organisational structure. This network structure that came to the fore was very agile and incredibly able to pursue small market opportunities, which would have been of no interest to larger companies (because either too small to be perceived, or too fast to register on the retina of large companies' eyes). These two features, the capability to exploit tiny market opportunities and the organisational agility reinforced the trend toward continual Schumpeterian innovation.

The disintegration of the large organisations led to a set of unintended consequences one of which was the emergence of the cluster form. The crisis put Meadow against a bifurcation, from which at least two alternative but internally coherent avenues departed. Had Meadow responded to the crisis following the road of layoffs and consolidation, the

---

[45] (Christensen 1997) has coined the term *value network* to refer to the set of environmental constraints that influences the agents' decision making. Although in theory the situation of the worker had not changed much, in practice the value network of the newly formed quasi-company was very different, and did consequently evolve along very different lines than it would have done had the worker retained his position within the integrated company.

cluster would have never emerged. Instead the history of Meadow together with the disintegration mechanism seen before caused the appearance of the cluster as a meta-organisation, which emerged out of the micro-actions of players who were pursuing their own individual fitness.

1.4 THE RISE OF BUSINESS ARCHITECTS

The increase in internal diversity got a further boost from the improved ability to scan the environment, thanks to the multiplication of the number of actors and to the refining of their perceptual capability, both caused by the phase specialisation mechanism. The secret to the exploitation of new opportunities came from the capability to quickly reconfigure the supply chain or value chain of phase specialised activities. Because the actors had a history of collaboration, an excellent knowledge of each others' capabilities and a well-tested trust relationship, the setting up of new networks within the cluster could be done with relative ease.

Many new product development projects turned into jigsaw exercises and new figures emerged to co-ordinate these temporary networks. One can easily imagine how the process of vertical disintegration, when pushed to its extreme consequences, would leave actors endowed with knowledge of the network but with no tangible assets. It is also easy to imagine that extreme vertical disintegration would affect first firms at the end of the value chain, i.e. firms with knowledge of the final markets. With the increasing complexity of the industrial and social structure in Meadow, some particular agents focused more on the interface with the distributors and final users, thereby shifting their actions from doing things to co-ordinating actors. The figures that emerged were business architects, people that could match external requirements for new products to internal offerings, and could lead the reconfiguration of new internal *ad hoc* networks to satisfy those needs[46]. Sitting at the end of the value chain they could scan the changes in demand and filter down the information to the supply chain[47]. The advantages of this

---

[46] These figures are real. In Prato they are called *Impannatori* ((Becattini 1997), (Malone and Laubacher 1998))

[47] Ronald Burt (Burt 1992) has coined the term *structural hole* to describe the highly favourable position of sitting between two networks (in our case, the external world of demand and the internal system of production) and mediating the flow of information between them.

position were speed and distributed leadership. With demand becoming more complex, the need to satisfy the vagaries of the demand of sophisticated types of yoghurt or fancy aromatised butter required essentially speed. This was achieved by coordinating the emerging project-based structure of the cluster. With the increase in the export activities, the increased distance with customer base required a specialised interface that could deal with the interpretation of demand, transfer the often tacitly expressed needs into a set of specifications to be passed to the productive structure, select and co-ordinate the right configuration for the specific order, and finally take care of delivery. The business architects became the pivotal coordination elements of the system, providing a crucial channel of communication between the external demand and the internal structure.

1.5 BASE MULTIPLIER AS AN AUTOCATALYTIC SET

It is well known in economics that a transaction can have a cascading effect on successive transactions. An export transaction can feed back into successive tiers of the supply chain creating an aggregated income for the cluster, which could be orders of magnitude bigger than the amount of the initial export. If X represents the income of an export activity, let's suppose that a fraction of it is spent in the local supply chain. This may lead to a second wave of transaction, then to a third and so on. If we call α the fraction transmitted along the supply chain, it is easy to see that the higher α[48], the higher the size of the aggregated income. Key to the ability of the multiplier to deliver its benefit is the concentration of activities behind the final export to be concentrated into a geographic area. This observation leads immediately to a non-linear relationship between multiplier and aggregated income. In fact, the positive correlation between the value of α and the concentration of economic activities within the cluster area indicates that the more the cluster internalises the socio-economic activities that are related to the initial

---

[48] the formula for the aggregated income Y is

$$Y = X + \alpha X + \alpha^2 X + \alpha^3 X + \ldots + \alpha^n X$$

which for large n can be approximated with:

$$Y = X * \frac{1}{1-\alpha}$$

See Fujita et Al. ((Fujita, Venables et al. 1999)) for more details export transaction and the more the cluster can diversify within those activities, the higher will be α and the final aggregated income. The multiplier formula captures the fact that as the size of the cluster economy grows, it becomes more convenient to produce a wider distribution of more sophisticated products and services within the cluster, thus giving rise to an increasing returns dynamics. Now if the cluster's income depends on the length of the supply chain then the higher the number of transactions performed in the supply chain, the higher the aggregated income. As a consequence of the previous point, the aggregated income will depend on the mechanisms that make the supply chains thick, complex and localised. What are they?

The disintegration of the large organisations into networks of interactive autonomous units and the centripetal attraction of location attractiveness (discussed earlier) caused a net increase in the firm population in Meadow. Flexible specialisation[49] together with Shumpeterian competition acted to make the supply chain composed of fine-grained units by creating new branches in the supply chain tree. Taken all together these forces made the supply chains in Meadow more similar to the delta of a river than to a set of linear chains. The net effect was to strengthen the action of the base multiplier, which, in its turn, by increasing the aggregated income of the cluster, reinforced the conditions for further complexification of the supply chains. The base multiplier captures the fundamental logic of clusterised economies, by which seemingly unrelated mechanisms (disintegration, specialisation, location attractiveness, ratio between neoclassical and Shumpeterian competition, etc.) interact as an autocatalytic set to promote the formation of a complex entity, the cluster, able to optimise exploration, by maximising internal diversity and connectivity.

1.6 NETWORK-BASED INNOVATION PROJECT

The network structure that emerged after the crisis led to the emergence of another practice in Meadow: network-based innovation. An example of this was the ice project. Although export had been one of the reasons of the prosperity, only few of the local products were exported. Most were consumed locally due to their short shelf life. A

---

[49] For more details on flexible specialisation see (Piore and Sabel 1984) and section 2.5.5

group of people that used to meet in a local pub[50] noticed that there was a huge reward waiting for those entrepreneurs that could solve the problem of export of short life products. They decided to tackle the problem and started meeting after work. The first research avenue that was tried was to transform short life into long life products, as it had been already done with milk and cheese. The alternative idea was to try to provide an environment during transportation that made short life products last longer. The solution was well known: ice.

This project represented a turning point for the system because its success legitimised the network approach to innovation. When the group within Meadow achieved its objective, it had no particular reason to stick together because in a project-based organisation profit accrued to the individual firms and not to the network. The disappearance of the group did not cause any visible change in the structure of the cluster, but left instead two intangible legacies. First, organising by project acquired the status of success and changed the local culture and second, it left in place a privileged series of information channels between the organisations and people that had been part of the ice project (**Figure 15**).

[50] In Silicon Valley the mechanism was similar: "*Every year there was some place, the Wagon Wheel, Chez Yvonne, Rickey's, the Roundhouse, where members of this esoteric fraternity, the young men and women of the semiconductor industry, would head after work to have a drink and gossip and brag and trade war stories about phase jitters, phantom circuits, bubble memories, pulse trains, bounceless contacts, burst modes, leapfrog tests, p-n junctions, sleeping sickness modes, slow-death episodes, RAMs, NAKs, PCMs, PROMs, PROM blowers, PROM blasters, and teramagnitudes, meaning multiples of a million million*" (Saxenian 1994) p. 33

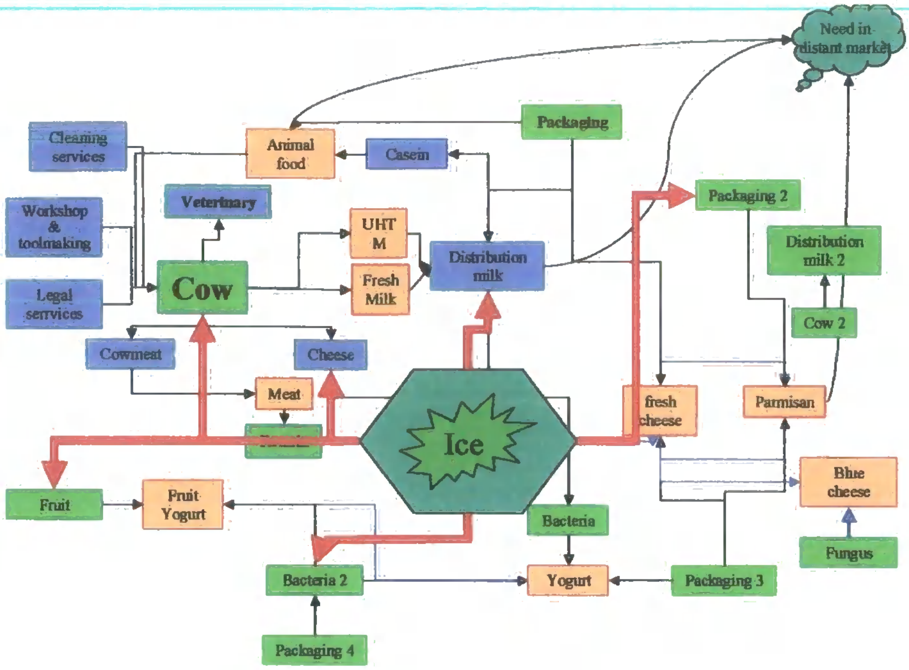


Figure 14 - Coevolution and project-centric organisations

The logic of network based innovation project was as follows. When economic demand transformed itself from a relatively regular succession of large homogeneous orders into a cascade of smaller and internally dissimilar orders, coming from a multitude of source, supply had to become agile. Being able to satisfy new niches required the capability of being constantly on an innovative edge. Large organisations were at a disadvantage, whereas systems of small organisations were better equipped, provided that mechanisms for co-ordination were in place. This is where organising by project kicked in and where Meadow's internal diversity of played a fundamental role. A cluster made of independent companies, characterised by a high internal diversity and connectivity, could innovate simply forming new sub-networks, whose structure (i.e. nodes and connections) was engineered to match the order's specifications. This means that groupings of firms, or sub-networks, formed and disappeared at each innovative order[51]. This fundamental change highlighted the transition from the firm as the unit of economic and managerial analysis, whose fundamental purpose is dominion of its market as a way to achieve survival, to a more complex environment in which the unit of the

[51] An analysis of this mode of organising is in (Arthur, DeFillippi et al. 2001)) and (Storper 1997)

analysis became the "limited life"[52] project, that is, the grouping of firms, whose purpose instead is dissolution

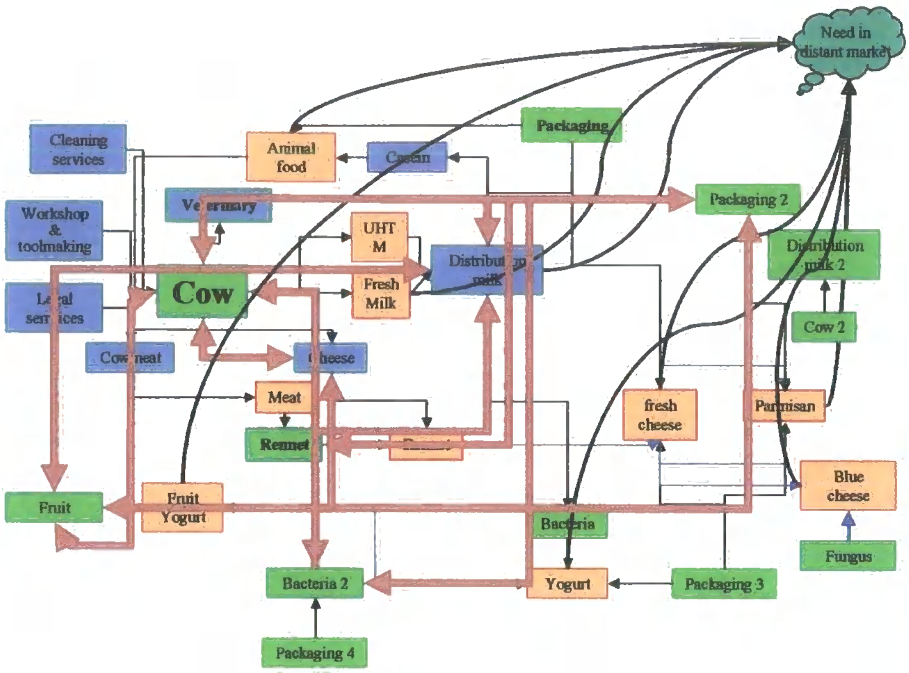


Figure 15 - Coevolution and project-centric organisations

1.7 CONSTRAINTS AND COEVOLUTION

I have summarised in table 1 the main innovations and transformations that took place in Meadow. The history of Meadow was shaped by a series of innovations, which can roughly be divided into two periods. The innovations/discoveries of the first period required a concourse of serendipitous circumstances to take place, exhibited a radical nature, were context-independent and helped to shape the main technological trajectories, business sectors and behavioural routines, which were followed in the rest of Meadow history. The latter period was characterised by context-dependent innovations with marked incremental, market-pulled, modular and organisational aspects.

Why were the two periods so different? Table 1 suggests an answer. The innovations of the former period required few or no pre-conditions to occur. Instead in the latter period each innovation required a set of culturally and historically developed

[52] The term is borrowed from (Hall 1999)

constraints[53]. These constraints were not generic, but instead highly specific to Meadow and its network culture. The relationship between innovations and constraints was akin to a coevolutionary dance in which the constraints shaped the basin in which the innovations took place and the innovations were used to generate more sophisticated and specific constraints. These newly developed constraints were used in order to improve environmental sense making and scanning to facilitate further innovations. The innovation process became more and more shaped by the context and specifically dependent upon the amount of diversity and the type of connectivity existing in the system. Diversity and connectivity had a profound impact onto the major parameters of Meadow, setting a trajectory of development based on the modularisation of production structures under conditions of high market uncertainty. The modularisation was achieved (and maintained) by pushing forward the boundary of task and work division via product and process innovation, thereby creating, by exploring smaller and smaller niches, further diversity and connectivity.

When an aggregate of organisations becomes locked into a self sustaining web of interdependent transactions and innovations, the system may undergo a phase transition, whereby the appearance of systemic properties represents the tangible sign of the emergence of a higher level of organisation, i.e. the cluster. The creation of the system Meadow was the result of the self-organisation of the parts within Meadow. Coevolution, that is the capability of the parts to generate the environment that constituted the context for the agents' actions, shifted the paradigm from the linear Newtonian logic, where causes and effects were clearly discernible to a circular logic, where causes and effects were circularly linked. It is as if the evolutionary process was happening on a rubber landscape where each step taken by any agent modified the environment of the others[54].

---

[53] As pre-conditions act as a relationship between two entities which condition their freedom, they can be interpreted as constraints. These ideas on the generative power of enabling constraints are presented in (Juarrero 1999) and (Kauffman 2000)

[54] The metaphor is by Kauffman in At Home in the Universe ((Kauffman 1995))

1.8 EXTERNAL SHOCK[55]

There were several risks for the survival of the original organisational form matured in Meadow:

- o Market stability: paradoxically a long period of market stability, in terms of customer buying patterns and/or long period of incremental innovation, may favour vertical integration. This makes economic sense as vertically integrated forms capitalise on economies of scale and learning curve effects;
- o External shock: changes in the external regulatory, fiscal and cultural environment.

Meadow was not an isolated place but it was part of a larger nation state in which an industrial revolution was taking place. The turmoil of change affected also the industrial policy of the young state. Until that point local systems had been free to explore their own 'existential' experiments in relative freedom. But, under the influence of the newly developing economic science, based on the principles of economies of scale, standardisation and scientific knowledge, new measures were introduced. These covered fiscal incentives, labour legislation, educational systems and communication systems. The objectives were manifold:

1. To favour the aggregation of the industrial base;
2. To encourage the evolution from a craft-based to a mass-manufacturing based industrial model;
3. To encourage the allocation of funding for research and investment to large industrial corporations;
4. To promote innovation based on accepted scientific knowledge.

All these effects contributed to changing the construction of the constraints/atmosphere that were at the base of Meadow's success. Especially the dominant culture, based on the opposition of the big firm as a superior vehicle for production against the small, did affect the way of thinking within our cluster. The asymmetry principle (so dominant in those times) stated that a large company could do

---

[55] The inspiration for this section comes from the historical analysis of the decadence of the French silk cluster in Lyon at the end of the 19th century. See (Piore and Sabel 1984)

everything that a small firm can do, but the opposite was not true. This principle based on the much revered concepts of economies of scale, learning curve effects, monopolistic control of market, influence of intellectual property rights issues and standards setting skewed the industrial policy agenda toward the principle that big is automatically good.

That same culture that did not see the network of tight collaboration and intense competition of the aggregate of small firms within the geographically bounded territory was then able (together with the formal legal, procedural and fiscal measures) to affect the delicate balance of the relationship of collaboration/competition within Meadow. The type of attack brought against the Meadowian way of doing business was qualitatively different from a contraction in demand, a recession, a change in customers' buying behaviours or other market fluctuations. It was not something against which the mechanisms of reconfiguration of internal networks could be used. The attack was at the culture of interactions of the cluster. The cluster had no antibodies against the cultural change imposed centrally because the clusters' agents were not aware of having developed an original form of production and innovation. As the history of industrial clusters demonstrate[56], the emergence of the cluster as a meta-organisation is not perceived for a long time by the clusters' agents[57], who still think that their behaviour makes sense according to a traditional reading of the situation. The change in paradigm promoted by the central state affected the very roots of the cluster that is the transaction style within the community.

But once the delicate mechanisms of vertical disintegration, phase specialisation, project-based innovation and production, tacit knowledge specialisation and transmission came under the combined attack of the economic science, industrial policy and fiscal

---

[56] For example the sentences by Saxenian ((Saxenian 1994) p. 164) that, *"although Silicon Valley's success has been based on collaborative practices, the region has long been dominated by the language of individual achievement"* and by Tom Hayes, founder of Joint Venture, that: *"our aim is to build a comparative advantage for Silicon Valley by building a collaborative advantage ... to transform Silicon Valley from a valley of entrepreneurs into an entrepreneurial valley"*, miss the point that the emergence of a complex place does not require an intentional approach by the agents to do so. There is no contradiction between the emergence of a complex place based on co-operation/competition and a language of individual achievement, if one considers that the two aspects refer to different hierarchical levels: the former to the systemic property of the cluster and the latter to the local behaviour of the agents. A complex reading of Silicon Valley (or of any cluster that happens to be in a valley) as a cluster would reveal at the same time the presence of the valley of entrepreneurs and the entrepreneurial valley as co-existing and originating one from the other without any contradiction.

[57] See Becattini's history of Prato about this point (Becattini 1997)

regulation, the distributed nature of the cluster collapsed. The paradox at the base of the cluster, whereby profits accrue to the individual firms, but the performance is a collective property of the web of trusted relationships, could no longer be sustained and imploded. As it was predictable Cow regained control of distribution; Cheese, Fungus and Yoghurt did eventually merge under the push of fiscal incentives (**Figure 16**). As a result of the increase in vertical integration and size, the firms lost in agility and increased their overheads. As a consequence, they started to reject small orders and to focus on the large ones. This increased the speed of migration toward mass market and bulk production. The leading innovative edge that had almost been an automatic consequence of the network structure of the cluster was consequently affected. Centralisation had other consequences too:

- The pressure of standardisation resulted in a pressure for codification of knowledge. The craftsman's know-how, effective but mysterious knowledge, that could not produce twice exactly the same object and that worked without full specifications, came to be regarded as antiquated and went out of fashion;
- The consequent shift of power from the distributed base to the technical elite further eroded the confidence in the Meadow's way of doing business.
- The new emphasis on centralisation and explicit knowledge did change innovation from distributed, architectural, often disruptive toward more predictable, centralised and less risky incremental innovation.

What is surprising is the relative rapidity with which the features of the mass-manufacturing model did invade Meadow.

A possible explanation has to do with the same economies of diversity whose effects were displayed during the phase of expansion of the network model. Diversity works in an autocatalytic fashion to generate more diversity via innovation. This process may become strongly non linear once a critical dimension of diversity is achieved. After that threshold, the system develops along a fast trajectory, where an increase in diversity triggers more innovation and more innovation triggers the start-up of more firms and a larger distribution of products. There is basically a multiplier in action.

But the same multiplier can work in the opposite direction, too. If changes in the regulatory and cultural environment turn the direction of variation of diversity from positive to negative, then the change in the system can be very rapid and the system turns from a network model into a more vertically integrated one.

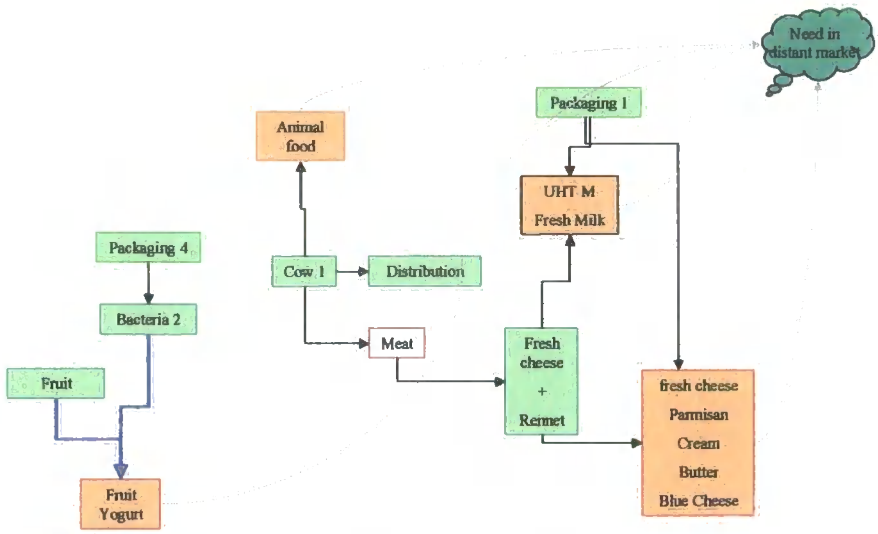


Figure 16 - External shock: imposition of mass manufacturing model

1.9 CONCLUSIONS

The dynamic introduced in the previous pages is counter-intuitive. What I've tried to show in the story of Meadow is that the formation of a highly dynamic but also ordered network is a result of an (largely, though not exclusively) endogenous process, whereby some set of cumulative mechanisms based on increasing returns dynamics coalesce and form the backbone of an idiosyncratic social and economic system. The story illustrates the power of co-evolution to speed up[58] the process of knowledge creation and innovation in the context of what I defined as economies of diversity.

Contrarily to the predicaments of classical economics, the system that emerges from this chapter is more similar to a self-sustaining vortex, which feeds its non-equilibrium situation by engaging into innovation at the agents' and at the system's level. It is this feature of industrial clusters that make industrial clusters an ideal terrain of 'experimentation' for complexity theory.

[58] The expression is by (McKelvey 2001)

The application of complexity theory to support the more traditional instruments of inquiry of industrial agglomerations, such as economic geography, economic sociology, and innovation theories, has helped me to generate a number of propositions regarding the dynamic properties of industrial clusters, which I will try to demonstrate during the rest of this thesis.

1. Industrial clusters are complex systems at *the edge of chaos*. I will propose a method to distinguish self-organising from non self-organising agglomerations, by making use of first, the distinction between aggregates of parts and systems, and second, power law theories.

2. Two of the critical parameters to make sense of clusters and distinguish them from other types of industrial agglomerations are their degree of internal diversity and their connectivity. Diversity and connectivity have a profound impact onto the major parameters of industrial clusters. Clusters are socio-economic entities that excel at modularising their production structure under conditions of high market uncertainty. The modularisation is achieved (and maintained) by pushing forward the boundary of task and work division, thereby creating, by exploring smaller and smaller niches, further diversity and connectivity.

3. Clusters seem to explore the space of diversity within the envelope of their extended value chain at a higher rate than non-clusters.

Innovation	Pre-Conditions	Type of Innovation	Technological Innovation	Product Innovation	Organisational Innovation
Discovery of UHT		Radical Technology -push	UHT	UHT milk	Cow Emergence of phases
Discovery of Cheese making process	Complementarity between meat and milk production Funding	Radical Technology -push	Rennet	Cheese	Spin off of Distribution and Rennet Partnership
Discovery of long life cheese	Partnership Experience with similar Innovation (UHT)	Incremental Market-pull	Exsiccation	Parmisan	Loops Increase in coordination (interactive production chains)
Discovery of casein	Availability of scientific knowledge Funding	Radical	Casein	Animal food	Autocatalytic loop
Diffusion of yoghurt	Location attractiveness Technology transfer Sophisticated market Funding	Micro-radical Technology -push	Yoghurt bacteria	Yoghurt	
Discovery of blue cheese and fruit yoghurt	Location attractiveness Technology transfer Sophisticated market Funding Shumpeterian competition Price competition Product diversity	Modular Market-pull		Blue cheese Fruit Yoghurt	
Modularisation and unbundling of integrated organisations	Location attractiveness Shumpeterian competition Price competition Sophisticated supply chain Flexible production systems Availability of micro or social funding Social capital: network of relationships and trust Product diversity	Business model Innovation			Disintegration and Increase in phase specialisation
Project based organisation	Shumpeterian competition Price competition Sophisticated supply chain Social capital: network of relationships and trust Knowledge diversity Modularity of competencies Network business architects Culture of knowledge sharing	Business model Innovation			Network based innovation Project based networks

Table 1: Evolution of cluster form and types of innovations

Chapter 2: Literature Review

2.1 INTRODUCTION

From the almost accidental rediscovery[59] in Italy of clusters nearly 30 years ago, the general topic of organisations (or groups of organisations) structured as a network has steadily increased in importance and probably now represents, also thanks to the Internet, one of the most important areas of management and business studies[60]. Networks constitute the context of this thesis and are explored from the point of view of a specific type of network, the industrial cluster.

Making sense of networks requires a theory able to explain network's distributed nature. Complexity theory constitutes one of the latest approaches to networks. Complexity theory is really an umbrella term for a diversified set of theories, frameworks and models that has grown to incorporate systems theory, chaos theory[61], synergetic, dissipative systems theory[62] and the contributions from the SantaFe Institute[63] in New Mexico. In essence complexity is a theory of networks.

Perhaps the most ambitious claim by complexity theorists concerns the universality of their models. The dynamic descriptions of systems, they claim, transcend the specifics of the individual networks and apply from biology to cities, from chemistry to economics. Krugman in "The self organizing economy"[64] supports that claim, showing that economic system, specifically towns, behave in a way known as self-organised criticality. Examples of self-organised criticality are nearly endless[65]. Self-criticality is not the only example of model that apply across very different types of networks (though it is the best):

---

[59] Marshall was the first to describe and formulate an economic theory about industrial clusters at the end of the 19th century (Marshall 1919)

[60] (Nohria and Eccles 1992)

[61] (Gleick 1987)

[62] (Nicolis and Prigogine 1989)

[63] (Waldrop 1992)

[64] (Krugman 1996)

[65] For a review see (Bak 1997; Buchanan 2000). This class of phenomena show a unique mathematical behaviour, the power law, which has become its identification tag. The power law seems to cross the boundaries of completely different systems revealing (still in an intuitive way) some hidden commonalties between them

coevolutive dynamics, self-organising properties of complex systems, emergent properties driven or not by simple rules, dynamics associated with critical amount of diversity and micro-diversity in the system, Bénard-like behaviours are common to many different systems. These are all dynamic patterns that show that, underlying the incredible variety of manifestations that testify of the inherent diversity in systems, there are also forces that transcend the context and nature of systems. These patterns emerge when two basic conditions are present: first, multiple channels of non-linear feedback exist between the constitutive elements of the system, and second, the system is in a dissipative or non-equilibrium state, that is, the system depends on the exchange of energy/information with its environment to survive.

The complexity approach poses a number of questions in terms of methodology, underlying philosophical approach and relationship with other fields of research. The literature review will explore some of these issues and describe the main streams of complexity theory. In particular the topic of self-organised criticality and other theories of networks will receive a particular attention, as they will be used as tools to demonstrate some of the main points of this thesis. But complexity is more than that: Castells defines it as "*a method for understanding diversity*"[66]. In the context of networks, complexity as a theory that focuses on diversity helps us make sense of the dynamic features of the *network organisation*. The issue of diversity is central to this thesis and therefore a large part of this literature review will be devoted to its description and analysis. Diversity is explored from two points of view. First[67], I revise the current interpretations of technological and social diversity as contributing to innovation, avoiding the trap of technological lock-in and acting as an *insurance policy*[68] against uncertainty and ignorance. Second, diversity, claims Kauffman[69], is a pre-requisite for self-organisation and emergence of progressively more complex entities (networks). These theories and models are very relevant for the interpretation of networks, but curiously, there are not many studies[70] that have actively taken this standpoint and applied it to industrial clusters.

---

[66](Castells 2000) p.74

[67] I have extensively used for this part the excellent text by Stirling (Stirling 1998)

[68] (Naeem 2002)

[69] (Kauffman 1995; Kauffman 2000)

[70] I am aware of the studies by (Grabher 2000; Allen 2001; Rullani 2001)

The other part of the literature review covers the field of industrial clusters. This is a very rich and highly diversified field. I introduce[71] the major schools of interpretation of the phenomenon of clusters. I have not covered the general topic of the network organisation, preferring to focus instead on the specific aspect of industrial clusters with a particular attention to the Italian examples.

At the end, what keeps everything together is the idea that clusters are a specific form of network that survive by continuously generating (and adapting to) innovations in a molecular and distributed form. The importance of the problem of industrial clusters lies exactly in this dimension.

Summing up the objectives of the literature review are:

- introducing the context of industrial clusters and placing it in the appropriate historical frame
- highlighting what has been done in the field and suggesting what needs to be done;
- suggesting important frameworks (such as complexity and the analysis of social and economic diversity) and variables for the empirical exploration of clusters;
- offering a new perspective on the topic of networks and clusters.

I will cover the following broad arguments:

2.2 Industrial clusters

2.3 Complexity theory

2.4 Power law phenomena

2.5 Diversity

The order of the main chapter reflects the following logic. First, I introduce the unit of analysis of this thesis, the industrial cluster, and present the main classical theories that have been put forward to explain the properties of clusters. Most of the interpretations of clusters make use of a variety of economic arguments such as economies of scale/scope, increasing returns, localisation and urbanisation economies, technological trajectories, etc. I will briefly introduce these theories and models in section 2.2. Although many scholars make an implicit use of concepts and ideas taken from complexity theory, there is very little evidence in the literature on industrial clusters of a systematic use of the central concepts of complexity theory[72]. I believe instead that self-organisations, emergence, autocatalytic sets, power laws, and diversity are fundamental to achieve a convincing explanations of industrial clusters. Hence the rest of the literature review is dedicated to the analysis of these concepts. I will start with an historical review of models of complexity theory. I will then move to the new area of power law and to the astonishing evidence that connect the emergence of power laws to phenomena as disparate as earthquakes, size of cities and supply chains. I will discuss the three dominant power law frameworks: rank-size rule[73], scale-free networks and self-organised criticality and argue that they are the manifestation (or an indicator) of a self-organising behaviour. The final part of this chapter will introduce the topic of diversity and connectivity. I will try to show, first, that diversity is a central, albeit neglected, topic in economic, innovation and cluster studies and second, that diversity is inextricably connected to the morphogenesis and survival of self-organising systems.

2.2 INDUSTRIAL CLUSTERS

2.2.1 SOME HISTORICAL ELEMENTS ON CLUSTERS

"*Throughout the nineteenth century, two forms of technological development were in collision. One was craft production. Its foundation was the idea that machines and processes could augment the craftsman's skills, allowing the worker to embody his or her knowledge in ever more varied products: the more flexible the machine, the more widely applicable the process, the more it expanded the craftsman's capacity for productive expression. The other form was mass production. Its guiding principle was that the cost of making any particular good could be dramatically reduced if only machinery could be substituted for the human skill needed to produce it. Its aim was to decompose every handwork task into simple steps, each of which could be performed faster and more accurately by a machine dedicated to that purpose than by a human hand.*"[74]

---

[71] According to Storper (Storper 1997)

[72] The only exception that known to me is (Rullani 2001). Rullani's paper uses complexity theory to re-interpret the concept of space in economic geography. Industrial clusters are the object of his analysis.

[73] Allen ?? p.27-28

[74] (Piore and Sabel 1984) p.19

This long passage articulates the basic idea of Piore and Sabel's seminal book[75]. The contrast between the two systems (craft production and mass manufacturing) revolves around the crucial relationship between the use of technology and the use of human resources. In the craft production system flexibility in the use of technology is seen as an augmentation of human creativity, and adaptability in the use of multipurpose technology is used to continuous adapt to customers' demand. This is achieved via batch production and specialisation of craftsmen, and requires a loosely connected network of producers. In the second half of the 19th century this form of production was marginalized and mass manufacturing took the lead.

The victory of the mass production paradigm in Piore and Sabel's account was never complete and, on the contrary, represented only one side of the coin of the so-called *industrial dualism*[76]. Mass production consisted in the production of general purpose good (therefore able to serve a mass demand) by means of specialised machinery. Yet this specialised machinery could not be mass-produced due to its small market size. Therefore its production was carried out by smaller organisations, which responded not to a mass market paradigm but to a flexible paradigm, whereby skilled workers used general purpose machinery as an extension of their skills to produce in a batch mode a few items of the specialised machinery required by the mass manufacturers. According to Piore and Sabel, the triumph of mass-market paradigm could not do without the parallel flexible, craftsman-based segment. It is useful at this point to clarify better the different meaning that the term specialisation acquires in the two paradigms. In the mass-market paradigm, specialisation is acquired by a process of labour division. Using the words of Adam Smith:

"In the lone houses and very small villages which are scattered about in so desert a country as the Highlands of Scotland, every farmer must be butcher, baker, and brewer for his own family. [...] The scattered families that live at eight or ten miles distance [...] must learn to perform themselves a great number of little pieces of work, for which, in more populous countries, they would call in the assistance from those workmen"[77]

and

[75] (Piore and Sabel 1984)

[76] (Piore 1980)

[77] (Smith, 1776)1979: 122) cited in (Maskell 2001) p.221

"In a tribe of hunters and shepherds a particular person makes bows and arrows, with more readiness and dexterity than any other. He frequently exchanges them for cattle or for venison with his companions; and he finds at last that he can in this manner get more cattle and venison than if he himself went to the field to catch them. From a regard to his own interest therefore, the making of bows and arrows grows to be his chief business, and he becomes a sort of armourer"[78]

Specialisation is therefore a by-product of the process of labour division. The productivity increase, resulting from partitioning a task into more minute sub-tasks, bears the price of increased rigidity and carries the necessity of devising mechanisms for coordination. The use of technology (i.e. more progressively sophisticated machinery) is seen as a way to increase productivity via the labour division mechanism. This vision of economic development was shared both by Adam Smith and Karl Marx[79]. However this increased path toward division of labour could not occur without an expansion in demand. This dependence on the size of market demand ensured that labour division, increased productivity and heightened specialisation could only happen in a mass market.

There is however a different meaning to specialisation that implies a different relationship with technology of which Marx was fully aware: *"in handcraft and manufacture the workman makes use of a tool, in the factory the machine makes use of him"*[80]. The workman that makes use of a tool, does so in order to improve versatility and product quality. In this case the machine increases at the same time the worker's specialisation (because it allows the production of new or improved things) and flexibility (as the use of technology together with an increase in craftsman's knowledge extends the potential range of use of craftsman's skills). Specialisation results from an increase in knowledge not from the deskilling that Marx lamented. This second type of specialisation is necessary in order to understand the emergence and the historical persistence of distributed forms of production and innovation such as the industrial clusters.

As Piore and Sabel rightly point out mass manufacturing requires a mass demand. By conversion when markets are too small or subjected to high levels of fluctuations, mass production becomes inefficient. In these markets the flexibility of the single craftsmen together with the integration of their specialisations in the context of a social community

---

[78] (Smith (1776)1979: 119) cited in (Maskell 2001) p.219

[79] (Piore and Sabel 1984) p.25

[80] (Marx 1967) p.422

made possible a continuous level of innovation and a permanent state of adaptation to the vagaries of the market fringes. Some historical examples will provide some evidence to this point.

In the nineteenth century several centres[81] became symbols of distributed systems of production: "*silk in Lyon; ribbons, hardware, and specialty steel in neighbouring Saint-Etienne; edge tools, cutlery, and specialty steel in Solingen, Remscheid and Sheffield; calicoes in Alsace; woollen and cotton textiles in Roubaix; cotton goods in Philadelphia and Pawtucket, ...*"[82] and the list could go on. Piore and Sabel identify three historical types of flexible production systems: municipalism, welfare capitalism or paternalism and familialism. All three types share some basic features. First, the economic activity was not separated from the territorially bounded social community in which it took place. This means, as for modern clusters, that economic interpretative frameworks must be supported and integrated with sociological[83] and historical analysis. Second, the relationship with the market is strongly dynamic. The cluster (or district to use Marshall terminology[84]) operates at the fringes of mass markets manufacturing a wide variety of products for demanding and differentiated market niches, some local but most international. Third, flexibility in the use of specialised resources and tools (or machinery) serves to augment the versatility and quality of craftsmen's actions. In terms of technology, flexibility meant a) frequent changes of input materials and techniques either to satisfy shifts in market demand or to anticipate markets[85] and b) recurrent changes in the production of products in a batch mode within a general family of products. Fourth and finally, districts elaborated institutional structures (through local politics) that mediated the conflicts between social actors and built behavioural patterns that allowed cooperation to coexist with competition.

---

[81] (Sabel and Zeitlin 1997)

[82] (Piore and Sabel 1984) p.28

[83] this point is at the base of modern sociology of economics, see Granovetter (Granovetter 1985), (Burt 1992), (Coleman 1990), (Bagnasco 2001), (Putnam 1993), (Fukuyama 1995).

[84] (Marshall 1919)

[85] An example of flexible machinery is the Jacquard loom. This is sometimes referred as a proto numeric controlled machine as punched cards were used to change the patterns of the goods produced. The Jacquard loom reduced the setup time between different batches, reduced the minimum size batch cost and required a new set of skills from the operators, thereby increasing their specialisation. The

Municipalism was the most important of the three. The 19th century silk Lyonese cluster was the best example of this form of organising production and innovation, defined by Frederic Le Play as *fabrique collectives*[86]. The fabriques collectives were like an inverted funnel where a large body of small and micro firms shared the same premises often around some production facilities. Typically fabriques collectives were diffused when barriers to entry (low capital intensity) were small and knowledge was specialised but diffused in the local territory. This confederation of small shops[87] were each firm specialised in a different production phase and "*before the widespread use of electric motors, these shops might be grouped in large buildings that housed a steam engine and a system of belts; the belts transmitted torque to workrooms that could be rented by the day*"[88]. The overall system was coordinated by a final organisation[89], either a large company or a merchant house with manufacturing knowledge and operations. This organisation was in charge of product commercialisation, provided often credit and raw materials and sometimes the final product assembling. As in modern clusters, the network of relationships shifted continuously according to the new styles required either by the market or by the end of the supply chain (trying to anticipate the market). The coordination of this network required a complex set of transactions that occurred horizontally, vertically and diagonally along the supply chain. Most of the transactions were dealt with informally, with trust and proximity used as mechanisms to punish free-riding behaviours and *intuitus personae*[90] being the crucial elements that lubricated the transactional system. On top of the transactional world, there were institutions which provided a) protection from the external world by means of unemployment benefits and credit facilitation in case of recession b) regulation of internal social conflict via fixing

---

Jacquard loom was developed in Lyon providing a strong indication that flexibility and innovation go hand in hand. (Piore and Sabel 1984), (Mokyr 1990)

[86] note 31, chapt. 2 (Piore and Sabel 1984)

[87] (Piore and Sabel 1984) p.32

[88] (Piore and Sabel 1984) p.32

[89] the coordinating organisation will evolve in modern version of clusters into what I defined as business architects. See section 3.2.2.4

[90] Judgment based on local tacit knowledge and on a set of weak links (to use Granovetter's parlance). The latter provided both an efficient mechanism of reputation verification and at the same time, in case of defection from the locally accepted standard rules of social/economic behaviour, a mechanism of exclusion by the benefits of the collective enterprise (see (Axelrod 1984) and (Axelrod 1997) for a detailed analysis of this point)

wages, work hours, etc. c) specialised training/education by setting up vocational schools and d) brand policy and quality control.

In welfare capitalism or paternalism a final firm provided the use of expensive production facilities, which were beyond the purchasing power of small artisanal firms, and outsourced whole phases of production to these firms. Examples of paternalism include the famous Calico printing firms at Mulhouse[91]. The leading firms organised a complex network of small firms via a mixture of hierarchy and community tools and incentives. For instance the leading firms founded vocational schools, research laboratories and a network of social institutions aimed at forming a community in which social and economic aspects were co-dependent. The main difference with municipalism consisted in the organising role of the leading firm.

Familialism is the third organisational flexible form reported by Piore and Sabel. This system was apparently invented by Alfred Motte in France around the 1850s as a response to a failed competitive strategy against well-entrenched textile mass producers. Motte devised a system whereby any time a new phase of production was needed a new company was formed by pairing a member of the Motte family with an experienced technician and providing them with the capital and resources necessary for succeeding. The newly formed company was independent and therefore found markets within and outside the family network. However, the familial bonding provided an easier transactional route and a higher cooperative mode for transactions internal to the family network.

2.2.2 MARSHALLIAN DISTRICTS

Marshall[92] noticed at the end of the 19th century localities where a high geographic and sectorial concentration of specialised small firms generated consistent and widespread economic externalities in spite of the lack of economies of scale. What type of externalities? Essentially three[93]: specialised labour pool, specialised network of local

---

[91] (Piore and Sabel 1984) p. 34

[92] (Marshall 1916)

[93] I am referring to classical economics externalities. Knowledge spillovers are sometimes classified as a fourth Marshallian externality

suppliers, common infrastructure and specialised business service organisations[94]. More in details the features of Marshallian districts (the terminology was coined by Marshall) are[95]:

- o Predominance of small firms over large
- o Closed system: most of supply chain transactions are intradistrict and most of significant suppliers are locally based
- o Long term commitment to suppliers and partners within districts
- o Diffused know how: "*The mysteries of the trade become no mysteries, but are as it were in the air*" due to high rate of knowledge *spillovers* among firms within districts
- o Alternative source of finance in terms of "*patient capital*" within districts
- o Workers' loyalty to community more than to single firms
- o Net positive rate of immigration compared to emigration

Marshallian districts capture well many of the features of the three historical cluster forms described above. However, modern districts have moved beyond Marshall analysis. They are the subject of next sections.

2.2.3 CLUSTERS TODAY

The clusters of the 19th century represented one of the possible evolutions of the craftsman-based production form. However the success of the mass manufacturing paradigm and the consequent paradigm shift from craftsmanship to mass manufacturing singled out, to the eyes of policy makers and industrialists alike, the 19th century cluster as an antiquated production form more akin to a residual form of organising work and production than mass manufacturing. The paradigm change together with active industrial policies favoured consolidation of existing business into larger entities, better able to exploit economies of scale and mechanisation. The result was the demise of 19th

---

94 (Gray, Golob et al. 1996; Markusen 1996), (Becattini 1990), (Sforzi 1990)

95 (Markusen 1996)

century industrial clusters. The story of the metalworking clusters in Sheffield, of the silk cluster in Lyon, etc., followed a similar path[96].

To confirm Piore and Sabel's analysis, clusters re-emerged after WW2 in those countries where a) the tradition of craftsmanship had been less affected by the dominant mass-manufacturing paradigm and b) the integration between societal and economical elements (typical of pre-industrial societies) had survived in communities well provided with social capital.

The literature on clusters is an extremely rich and ill-defined field of research that crosses several disciplines: economics[97], economic geography[98], management[99], economic sociology[100], technology and innovation[101], just to name a few. All this literature starts from the fundamental assumption that location should be an integral part of economic analysis. By necessity it has to cut across the disciplines mentioned above in order to make sense of a localised economic phenomenon, such as the cluster, in which purely economic reasoning fails to describe the interaction between location, work, production and social community. I will claim in this thesis that complexity theory needs to be added to the list above as theories based on the concepts of equilibrium and diminishing returns are unable to describe clusters as a self-generative phenomenon[102].

2.2.4 DEFINITION OF CLUSTERS

The current descriptions of geographic clusters focus on the properties of co-location or proximity, vertical disintegration leading to flexible specialisation, peculiar governance forms based on cooperation-competition mixture, presence of collective learning and

---

[96] (Piore and Sabel 1984) pp.35-37

[97] Marshall on spillovers and externalities, Hirshmann on backward and forward linkages (Hirschman and Otto 1968), Krugman on international trade and location (Kale, Singh et al. 2000), Arthur on increasing returns (Arthur), (Fujita, Venables et al. 1999)

[98] See Storper (Storper 1992), (Best 1990), (Amin and Thrift 1992), (Glaeser, Kallal et al. 1992), (Porter 1990), (Maskell 2001)

[99] (Porter 1998), (Arthur, DeFillippi et al. 2001), (Kogut 2000), (Powell 1990), the new literature on network and new economy (see (Malone and Laubacher 1998), (Evans and Wurster 1999), (Castells 2000))

[100] (Becattini 1990), (Brusco 1982), (Piore and Sabel 1984).

[101] (Dosi and Orsenigo 1985) on technological trajectories, (Lundvall 1992) and (Nelson 1993) on national innovation systems

diffused tacit knowledge, and economies of agglomeration. Different models based on an idiosyncratic mixture of all or some of the above features include: the neo-Marshallian or Italianate model[103], such as the textile cluster in Prato; high tech or 'hot spot' clusters (Pouder, 1996) such as Silicon Valley[104] or the Formula One cluster in the Oxfordshire[105]. Some authors extend the definition of cluster to embrace the 'hub & spoke' model[106], locally concentrated supply chain, the satellite industrial platform and the state centered districts[107].

The different forms of spatial aggregations reported above present few commonalties:

- aggregation over a geographically delimited territory;
- specialisation around a set of crucial designs, technologies and production techniques;
- presence of multiple forms of traded and untraded interdependencies;
- 'stickiness'[108] or 'embeddedness'[109]: the way in which practises of trade, production and provision of services are embedded in social systems and history;
- "*neither market nor hierarchy*" governance form[110].

Defining clusters is difficult. Some authors do not subscribe to the idea of clusters[111], preferring instead to talk of agglomeration of small firms subject to a regime of loose

[102] (Andriani 2001), (Rullani 2001), (Allen 2001)

[103] (Piore and Sabel 1984)

[104] (Saxenian 1994)

[105] (Henry and Pinch 1997)

[106] (Gray, Golob et al. 1996)

[107] (Markusen 1996)

[108] (Markusen 1996)

[109] (Granovetter 1992)

[110] (Powell 1990)

[111] see (Amin and Robins 1990) for a radical criticism of the flexible specialisation and transaction costs analysis schools. Martin (Martin and Sunley 2002) attacks the whole concept of clusters, although his target seems to be only the Porterian approach to it

coordination, others instead extend the definition of industrial clusters to include dense supply chains[112].

Porter[113] offers the following definition: "*a cluster is a geographically proximate group of interconnected companies and associated institutions in a particular field, linked by commonalities and complementarities*"

A more precise definition comes from Pyke[114]: Clusters are:

"*...industrial systems ... composed of (generally) independent small firms, organised on a local or regional basis, belonging to the same industrial sector (including all upstream and downstream activities), the individual firms tend to specialise in a particular production phase, organised together and with local institutions, through relationships of both competition and co-operation*"

The definitions above indicate the complex nature of clusters. The identification of the parts that form a cluster require the analysis of a) relationships based on complementarity along the value chain both downstream and upstream, b) relationships of cooperation along a horizontal direction based on the use of similar distribution channels, specialised inputs, etc., and c) relationships with institutions providing specialised skills, training, funding, information, etc. Clearly simple agglomerations of similar firms do not constitute a cluster. Neither the presence of cooperation alone is *per se* an indicator of cluster dynamics. The fuzziness of the cluster concept and definition makes difficult the identification of the boundary of clusters. As Porter claims: "*drawing cluster boundaries is often a matter of degree, and involves a creative process informed by understanding the most important linkages and complementarities across industries and institutions to competition*"[115].

The variability of clusters extends to their specialisation sectors[116]: textile, fashion, food, computer hardware, software, mechanics, optoelectronics, etc. They are also

---

[112] (Markusen), (Gray, Golob et al. 1996), (Nishiguchi and Beaudet 1998) and (Piore and Sabel 1984) for Toyota City

[113] (Porter 1998) p.199

[114] Pyke, 1992; quoted from Curran, J and Blackburn, R., 1994. "Small firms and local economic networks: the death of the local economy?" Paul Chapman, London

[115] (Porter 1991) p. 202

[116] (Porter 1991)

particularly diffused in all main Western economies[117], in large as well as small countries. The examples of Italy[118], Germany, Denmark[119], US[120] and France[121] are particularly famous. A recent TBR report[122] explores the situation in the UK. Porter shows that industrial clusters are relevant in the US and provides several examples of clusters diffusion in emerging countries[123]. Some clusters are dominated by small firms (in Prato firms' employees average is 7[124] with no company bigger than 200 employees), others show a mix of large and small firms (Hollywood, Silicon Valley, the advertising cluster in London[125], or the textile cluster in Biella, Italy).

Clusters have curiously escaped the attention of scholars for a long time. With the exception of Marshall, the first formulation of the cluster idea was formulated at the beginning at the 60s by Becattini and his Florence school[126]. This is due to a complex of circumstances: first, the separation between sociology and economics[127] made difficult the integration of these disciplines in field research; second, the idea of clusters went against the accepted wisdom of scientific management and mass manufacturing; and finally, the fact that the boundaries of clusters included different industrial sectors, and did not respect geographic boundaries, diluted[128] the economic effects of clusters across official statistics and different industries[129].

---

[117] In Italy they account for a third of the overall export, thereby significantly contributing to the GDP (Becattini 1997) p.124-128

[118] (Piore and Sabel 1984; 1990), (Best 1990)

[119] (Maskell and Malmberg 1999)

[120] (Saxenian 1994), (Storper 1992)

[121] (Storper 1992)

[122] (Trends Business Research Consortium 2001)

[123] Porter argues that the limited diffusion of clusters in emerging countries is due to the precarious state of the main social and economic macro-conditions, such as education, advanced transportation and communication infrastructure, banking, research, etc., thereby limiting the competitiveness of agglomerations of small firms.

[124] See (Lorenzoni and Ornati 1988)

[125] (Grabher 2000)

[126] (Rullani 2001) points out that Becattini introduced the idea of neo-Marshallian districts as early as 1961

[127] See (Granovetter 1992)

[128] (Porter 1991) p. 204

[129] The following sentence by Saxenian on Silicon Valley's engineers highlights that the awareness of clusters was hidden even to the clusters' agents: "*paradoxically, however, while the region's*

In what follows I will describe the main schools that have caused the resurgence of the discussion on industrial clusters. It is important to notice that this literature overlaps with the literature on regional economies and that the distinction between the main schools is difficult to draw.

2.2.5 THE FLEXIBLE SPECIALISATION SCHOOL

The school that sprung in the 70s around the figures of Becattini, Brusco, Piore and Sabel and many others discovered that organisational forms that seemed residual to the eyes of neoclassical economists (and policy makers alike) hid instead a complex network dynamics that made them at the same time innovative, resilient and flexible. This school that has become known as *flexible specialisation school* put at the centre of its analysis a series of local industrial realities (of which the most famous are the woollen cluster of Prato[130] and the ceramic tile cluster of Sassuolo near Modena[131]) mainly in Central and Northern Italy, that have become a paradigmatic example of the so called neo-Marshallian, Italianate or Third Italy industrial district. The original points developed by the flexible specialisation schools concern a vast body of theories and rich empirical evidence in support of those theories. Adopting as unit of analysis a location made those economists and social scientists acutely aware that the tools of economic analysis in isolation were insufficient to crack the dynamics of locationally bounded systems of production, such as industrial clusters. Industrial clusters rely as much on economic drivers, profit, efficiency, scale economies, competition, etc., as on social drivers such as peer reputation, norms of reciprocity and system of relationships (familial or others). The mix of the two forms a community in which opportunistic behaviours (the free rider problem) and the type of behaviour characteristic of the "*tragedy of the commons*"[132] are avoided thanks to emerging coordination of the cluster's agents. Storper identifies three

---

engineers saw themselves as different from the rest of American business, they failed to recognise the importance of the networks they had created. Silicon Valley's entrepreneurs failed to recognise the connection between the institutions they had built and their commercial success. They saw themselves as the world did, as a new breed of technological pioneers, and they viewed their success as independent of the region and its relationships"(Saxenian 1994) p, 56). The history of Prato reveals even more paradoxical examples (Becattini 1997)

[130] (Piore and Sabel 1984; Becattini 1997)

[131] (Porter 1990), (Brusco 1982)

[132] (Hardin 1968), (Schelling 1978)

crucial contributions of the flexible specialisation school[133]. The first important contribution was then the realisation that economics, sociology and (to a lesser extent) anthropology were necessary ingredients for the comprehension of clusters.

This first insight led to the second: the mix of local conditions, deep history and economic interactions under the persistent survival of artisanal traditions could simply not converge to the optimising equilibrium, object of neoclassical economic analysis. The flexible specialisation school then posited that the evolution of a local system did not point toward a single *best way,* dictated by resource optimisation and economic equilibrium, but depended on local historical circumstances and unique histories. According to this view, clusters were not vestigial remainders of a pre-modern era or products of an incomplete process of industrialisation, destined to be reabsorbed into the standard Fordist model, but, on the contrary, they represented an alternative form of organising production around social communities. This implies that the direction of development is not necessarily from the small to the large and from the fragmented to the integrated, but could, on the contrary, be based on the exploration of the economic effects of network structure and phase specialisation.

The third important contribution, already mentioned above, focus on the simultaneous achievement of increasing specialisation and flexibility. In a Fordist paradigm, flexibility and specialisation are mutually exclusive. Specialisation entails an increase in the rigidity of the manufacturing process and therefore a loss of flexibility. What Piore and Sabel observed was that the demand for flexibility in the Post-Fordism era can be met by an increase in specialisation in a network context, in which the nodes undergo a constant process of reconfiguration to accommodate changes in market demand or in technological supply. With the risk of oversimplifying, we could say that specialisation is embedded in the nodes, flexibility in the shifting patterns of links. Piore and Sabel went on to apply the flexible specialisation model as a general blueprint for Post-Fordism, extending the lessons from geographic clusters to large firms and corporations.

The attention of the flexible specialisation school for decentralised forms of production and innovation contributed to the establishment of the network paradigm[134] in

---

[133] (Storper 1997)

[134] (Nohria and Eccles 1992)

management and economics. "*Anything that can be done in the vertical way can be done more cheaply by collections of specialist companies organised horizontally*"[135]. This sentence reveals how deeply has the network paradigm affected management thinking. The flexible specialisation school imposed to the scholars' attention the importance of socio-economic networks by a) providing empirical evidence of their existence and b) developing frameworks in which localisation, territory and disintegrated production (and resource ownership) are all connected via idiosyncratic learning processes. The territory (be it region, municipality or an intermediate form) becomes the place where specific forms of knowledge flows take place in a medium, which is not reducible either to hierarchy or to market dynamic[136]. The contribution of the flexible specialisation school was so incisive that scholars such as Michael Best[137] predicted that the future belonged either to the flexible Japanese company, arranged in the form of a pyramidal network, or to the Italian horizontal network district. The discovery of this new dimension of managerial and economic action, location specific, raised a whole new set of issues, related to why, how and where clusters form, what keep them together and what type of industrial policy promotes or favours cluster competitiveness. The flexible specialisation school, centred on the sociology of relationships in the economic environment of the community and on the role of institutions in promoting cooperation in the context of competitive relationships, could only partially address the new issues that the recognition of clusters had raised. Other schools and frameworks have explored some of these issues.

2.2.6 THE EXTERNAL ECONOMIES SCHOOL[138]

If the flexible specialisation school started form the empirical cases of the Italian district, the California school moved its first steps from the empirical analysis of the agglomerations in Silicon Valley[139], Hollywood[140], Los Angeles[141], New York[142] and many others.

---

[135] Andy Grove, chairman of Intel; quoted in (Saxenian 1994) p. 142

[136] (Powell 1990)

[137] (Best 1990)

[138] the definition is taken from Storper (Storper 1997)

[139] (Saxenian 1994)

The Hollywood case is paradigmatic. During the 50s the dominance of the 5 big Studios ended with the radical disintegration of the integrated Fordist model, that had until then been dominant. The process started with the outsourcing of the most risky and less profitable operations in response to two exogenous crises that hit the industry: the advent of the television and the antitrust action of the American government which forced the Studios to divest their control of theatres[143]. The loss of control of a captive market and the decrease in tickets sold due to the alternative offered by television forced the Studios to react with two strategies: on the one hand, they reduced their overhead by cutting back the number of movies produced; and, on the other hand, they frantically tried to increase flexibility to reduce costs and innovate to differentiate their product from television in order and win back customers. The strategy, although partly successful, started a cycle of unintended consequences that resulted in the disintegration of the industry. The search for innovation and flexibility required the increase in use of specialised inputs on a temporary basis. This caused an increase in the average cost of movie to which the industry reacted with further disintegration. One of the unintended consequences was the end of the star system, whereby the studios would invest in a (relatively large) number of young actors, securing their control by means of long-term contracts, and would reap the reward with the few successful ones. The same mechanism worked for the major phases of movie production: preproduction (script, selection of location and actors, etc), production (filming and scenography) and postproduction (processing, editing, sound track, marketing, etc.). In the attempt to reduce overhead and innovate, the main professional figures started to be recruited on a temporary basis. In both cases the industry had to face a dramatic cost increase due to a) the market costs of actors, directors, technical specialists, writers, etc. and b) the sheer amount of external transactions necessary to set up each production. The reaction of the industry was the agglomeration around a specific location (in particular Hollywood) and the elaboration of specific norms of governance that transformed the contractual norms that governed the previous 'mass-manufacturing' Golden Age. Agglomeration, project-based

---

[140] (Storper 1982)

[141] (Scott 1988)

[142] (Porter 1991)

[143] the 5 big studios owned the almost totality of theatres in the large cities: 70% of first run cinemas in cities with more than 100.000 inhabitants - (Storper 1997) p. 85

governance system and flexible organisational network were used to fight the increase in transaction costs and achieve innovation and flexibility.

In sum, the contribution of the Californian school consists in the elucidation of the above dynamics via the introduction of the transaction costs[144] argument. Contrarily to the Coase/Williamson transaction costs analysis[145], the increase in asset specificity does not lead to internalisation of activities but on the contrary leads to a self-sustaining process of disintegration and agglomeration. Geographic networks of small firms are formed when the process of vertical disintegration triggered by demand and/or supply uncertainty is met by a parallel process of agglomeration over a limited territory, that counterbalance the increase in transaction costs and allows the co-ordination of a supply network.

The Californian school then arrives to similar results as the flexible specialisation school but with some additional insights into the dynamic of networks: the formation and development of networks is no longer dependent upon the thick historical substrate that characterise the Italian (and European) clusters[146], but becomes instead dependent on some endogenous dynamics, which start at some branching point and develop into a fully fledged space and place of industrial governance transactions system.

2.2.7 The Technology Schools

Central to both flexible specialisation school and external economies school is technology. Already Piore and Sabel had identified flexible forms of production (and that included large corporations structured in a flexible way as well as industrial clusters) as appropriate forms able to adapt to uncertainty (market and technological) and survive to competition via constant change. Large part of the Becattini's and others' analysis demonstrates that Italian clusters thrive at the fringes of mass manufacturing by constantly shifting the terrain of competition towards new and more sophisticated needs, embodied in short life market niches. The network form of production is then really a

---

[144] (Coase 1937), (Williamson 1985)

[145] we could speculate that, with regard to transaction costs and the structure of the firm, the Coase/Williamson analysis and the Californian School analysis are like the two branches of a bifurcation

[146] exemplified by (Putnam 1993)

form of adaptation through constant innovation[147]. Yet innovation and new technologies represent, according to the flexible specialisation and transactional schools, an organic consequence of an institutional arrangement, the network form, and of its dynamic properties: labour division and transactional agglomeration. They are neither the engine of agglomeration itself, nor the main determinant of its success. The magic of clustering remains elusive.

A complementary approach is represented by a diversified set of schools, theories and frameworks, which put at the very centre of their analysis innovation and the dynamic features of technology development and diffusion. The explanation of the origin of clusters, such as Silicon Valley and Route 128, is now attributed directly to the properties of technology creation and diffusion. Going back to Shumpeterian cycles of creative destruction, some scholars[148] wondered whether agglomeration of companies followed the birth and development of new industrial sectors around some forms of radical innovation. The study of localised networks of economic activities extends then to the fields of innovation and diffusion of technologies, not as ancillary activities of the institutional organisational form – the network – but as the cause of agglomeration itself. The central problem turns therefore to: how does innovation happen? Why there? And, could the pattern of recursive innovation be created ex novo?

A simple way to address the problem was to study the successful examples of the famous innovation islands: the Silicon Valleys of the world. The search for critical success factors led to the identification of the crucial links (especially in US) between entrepreneurships and centres of creation of new knowledge, the universities. It seemed that a precondition for successful regions in high tech sectors was the intellectual power of highly creative universities. Silicon Valley sprouted from Stanford and Berkeley, Route 128 from MIT and Harvard[149]. Clustering comes then to depend on the dense web of linkages between creation of new science and technology in universities and research centres, and the entrepreneurial activities that sprout from technology transfer. The connection between research and entrepreneurship gave rise to the *technopole*[150] idea

---

[147] this point will be better developed in chapter 4

[148] (Markusen 1986)

[149] (Saxenian 1994)

[150] (Castells 2000)

and industrial policy, whereby a formal science/technology critical mass of research and development activities was supposed to generate a dense web of university-firm links via spillovers and knowledge transfers and therefore to give rise to agglomeration economies. The entrepreneurial activities, which would have followed from the application of the new science/technology generated in the universities, should have in turn generated additional funding for research and a market-led direction for new research. The two sides of the coin, research in formal science/engineering institutions and entrepreneurial firms, were expected to generate a virtual self-sustaining circle of research, applications of research and commercialisation of science-based products. However, most *technopoles* failed to become new Valleys, despite the amount of R&D carried out, probably because the link between R&D and economic clustering is not necessarily science-led. The virtuous circle between scientific/technological innovation and economic prosperity was still elusive. Even throwing into the balance mission-led R&D, such as space and military investments in research, the question why certain localities turn into engine of prosperity and others fail, was not answered. The basic idea that the origin of agglomeration depended upon some sort of centripetal force, (some scholars[151] claimed), that created a critical density of linkages between research/technology and commercial spillovers (and that after the initial trigger the cluster's dynamic would become self-sustaining) did not explain how the spillovers take place and why many clusters seemed to follow a different dynamics.

A more systematic school of thought gathered in the 80s and 90s around the GREMI[152], a group of regional economists[153]. Their central idea the '*milieu innovateur*' is defined by Castells [154] as

"a specific set of relationships of production and management, based on a social organization that by and large shares a work culture and instrumental goals aimed a generating new knowledge, new processes, and new products. Although the concept of milieu does not necessarily include a spatial dimension, I argue that in the case of information technology industries, at least in this century, spatial proximity is a necessary material condition for the existence of such milieux because of the nature of the interaction in the innovation process. What defines the specificity of a milieu of

---

[151] (Markusen and et al. 1991), (Markusen, Hall et al. 1986)

[152] Groupment de Recherche Européen sur les Milieux Innovateurs

[153] (Camagni 1991), (Aydalot and Keeble 1988), (Castells and Hall 1994)

[154] (Castells 2000) p.419

innovation is the capacity to generate synergy; that is, the added value resulting not from the cumulative effect of the elements present in the milieu but from their interaction. Milieux of innovation are the fundamental sources of innovation and of generation of value added in the process of industrial production in the Information Age".

The innovation milieu school starts from the observation that the distribution of innovative activities in the advanced countries is highly localised and few centres around the world are responsible for large part of science and technology production. Even the R&D activities of large multinational corporations are performed (usually) in the home country and are localised in few large labs[155]. The fact that "*technological innovation is not an isolated instance*"[156] leads the innovative milieu school to focus on the macro-properties of the milieu. A milieu exists because innovation is fundamentally a sticky process, which requires an institutional setting made of companies, institutions, rules of competition and cooperation, distribution and quality of skills and knowledge, in which the effects of new technology discovery and diffusion reinforce the organisational logic of the milieu, triggering in this way, further discovery and further innovation. This school stresses the idea that a cluster is a meta-organisational arrangement, whereby production, work and resources are geared towards innovation and that the relation between innovation and milieu is provided by the self-reinforcing nature of synergy between "*knowledge and information, directly related to industrial production and commercial application*"[157]. Callon[158] defines *techno-economic networks* as composed of three poles: first, the scientific pole, consisting of independent research centres, universities and industrial laboratories, in charge of producing new knowledge; second, the technical pole, *which conceives of, develops or transforms artefacts destined to serve specific purposes*; and third, the market pole, that is, the set of users in charge of formulating needs. When the three poles become a "*coordinated set of heterogeneous actors, who participate collectively in the conception, production and distribution/diffusion of procedures for producing goods and services*"[159], sharing the same territory, then a techno-economic network emerges.

---

155 (Pavitt and Patel 1991), (Dunning 1988)

156 (Castells 2000) p.35

157 (Castells 2000) p.67

158 (Callon 1991)

159 (Callon 1991) p.4

This approach puts together the attention to institutions of the flexible specialisation school with the focus on technological innovation and localised learning of the approaches described above. However "*the GREMI group has never been able to identify the economic logic by which milieux foster innovation. There is a circularity: innovation occurs because of a milieu, and a milieu is what exists in regions where there is innovation*"[160].

A fundamental milestone to this discussion was provided by a group of evolutionary economists[161], geographers[162] and scholars[163] of technological innovation. They started with the observation that the accretion and diffusion of technologies (the main engine of economic change) takes place in a fashion similar to biological evolution. According to them, the dynamics of innovation is strongly path-dependent with the set of previous choices and contexts setting boundaries and constraints on the future trajectory of the system. The evolution of the system, its trajectory, does not follow the general laws of neoclassical economics, whereby an optimum allocation of resources is achieved and the system evolves along lines which allow predictability of outcomes. These laws would make the evolution of the system largely independent from local history once equilibrium is achieved. Instead, the system's evolution turns out to be a relational property of the context, history, and decisions taken by agents internal to the system and agents external but coupled to the system's boundary. If the choices taken by agent A depend on the contemporary choices of B, C and D (and context) and B's decisions depend on A, C and D, and assuming the system is under conditions of radical uncertainty[164] or ignorance[165], whereby the number of possible outcomes are in practical terms indefinite, the evolution of the system becomes not predictable and strongly idiosyncratic. The change of metaphoric inspiration from physics to biological evolution transforms the object of study from science into an historical narrative, whereby the evolution of the system comes to depend on a web of interdependencies that define local and irreversible

---

[160] (Storper 1997) p.17

[161] (Nelson and Winter 1982), (Dosi and Orsenigo 1985; Dosi 1988)

[162] (Lundvall 1992)

[163] (Mokyr 1990)

[164] Knightian uncertainty (see (Stirling 1998))

[165] see Stirling at p.15 for a discussion on the difference between uncertainty, risk and ignorance

trajectories. Multiple trajectories become then possible in correspondence of the same initial set of variables.

The second important point comes from a non-orthodox reading[166] of evolutionary theory. Traditional Darwinism thinks that evolution is free in the exploration of change through mutation and natural selection given history and context. However, evidence from history of technology[167] and palaeontology[168] reveal that the direction of change be severely constrained along certain pathways. These are large envelopes within which technological (or biological) change takes place around a basic technological design[169] (or body plan in biology[170]). Change, technological or biological, consists in the exploration of the almost endless variety of functions, tools and adaptations that a basic design can give rise to. For instance the basic design of the steam engine has found applications in almost any aspect of energy conversion tools. However once the basic design is given, evolution is constrained and nature plays all sorts of tricks[171] to adapt the basic design to situations for which this is difficultly adaptable. The series of changes that a basic design undergoes from initial competition with other basic design to dominant design stage[172] to maturity and eventually senescence are described by Dosi[173] as a technological trajectory

This can be defined as a "*path of technological development, drawing on a given set of basic scientific principles and propelled by an internal dynamics of improving performance in terms of few crucial design criteria*"[174]. Stated another way, the interdependencies among agents' decision making during phases of radical innovation, can give rise to new technological trajectories, characterised by strong irreversibilities in

---

166 (Gould 1977; Gould 2000), (Eldredge and Tattersall 1982; Eldredge 1995)

167 See the final chapter of (Mokyr 1990)

168 (Gould and Eldredge 1977)

169(Gould 1997), (Anderson and Tushman 1997)

170 (Mayr 2001)

171 (Gould 1997) has proposed that the sixth Panda's thumb as an example of tricks necessary to overcome the fundamental limitations to adaptation that arise when some elements or the whole architecture of the body-plan hinders adaptation

172 (Abernathy and Utterback 1978), (Teece 1987)

173 (Dosi and [et al.] 1988)

174 (Ergas 1986) p.8

the way resources, technical knowledge and generic means of productions are put together to service the new technologies and resulting markets. But irreversibilities almost by definition show unique properties that tie them to the place where they have been generated. The diffusion of the techniques of production and related cultural aspects outside the initial network takes place slowly as it requires a complex set of changes that attain the technological, consumers' behaviour and organisation of production spheres. In sum, whereas a traditional Darwinist reading would suggest that technical change is free to explore change in any direction and rationalist approaches would suggest that, given a set of technical changes, different socio-economic systems could converge onto an optimum use of the new technologies (which, once given the local technical systems, would all converge to the same equilibrium situation), on the contrary, according to evolutionary economists, different socio-economic systems would develop different ways of using technologies along technological trajectories, often embedded in local systems.

There are two further aspects of technological trajectories important for our discussion. First, a technological trajectory establishes a new technological paradigm[175], that is, a set of behavioural, cognitive and perceptual rules, together with their underlying assumptions. The fact that some localities can be at the forefront of a new technological paradigm insulates them from other localities. If the conceptual and perceptual filters used to make sense of data and turn them into knowledge[176] depend on the state of development of a technological paradigm, then the reality perceived by places will be different[177]. In fact, localities involved in innovation will develop specific criteria for abstraction, codification and diffusion[178] of information. As evidence shows us that

---

[175] (Kuhn 1996)

[176] (Boisot 1998)

[177] *The history of astronomy provides many other examples of paradigm-induced changes in scientific perceptions. Can it conceivably be an accident, for example, that Western astronomers first saw change in the previously immutable heavens during the half century after Copernicus, new paradigm was first proposed? The Chinese, whose cosmological beliefs did not preclude celestial change, had recorded the appearance of many new stars in the heavens at a much earlier date. Also, without the aid of a telescope, the Chinese had systematically recorded the appearance of sunspots centuries before these were seen by Galileo and his contemporaries. ....*
The very ease and rapidity of with which astronomers saw new things when looking at old objects with old instruments may make us wish to say that, after Copernicus, astronomers lived in a different world. In any case, their research responded as though that were the case. (Kuhn 1996), pag. 117

[178] (Boisot 1998)

innovation is strongly territorially bounded, this implies that the more a territorial system is engaged in innovation, the more it internalises the knowledge necessary to make sense and describe the results of the innovations.

Second, under conditions of non-equilibrium (where changes in supply of technologies and in demand of goods/services are fast), the acquisition of dynamic capabilities (required to adapt and survive) rests on a set of traded and untraded interdependencies among the agents at the various level of the system, whose interaction are not a priori definable. "*Knowing how to do one thing is frequently consequent upon knowing how to do another*"[179]. Many, if not most, of the untraded interdependencies rely on knowledge only partially or not at all codifiable and on a set of individual skills and collective routines[180]. The interactions between the knowledge base, skills, routines[181] and social organisations (where knowledge, skills and routines are enacted) define a community (or a set of overlapping communities), formed by the agents in charge of transforming existing knowledge into innovations and new knowledge. Because the process of creation and diffusion of knowledge is inherently sticky, due to the tacitdness of new knowledge and the learning-by-doing aspect of skills and collective routines, the capability to understand and make use the new knowledge diffuses slowly outside the community.

To sum up: the reasons why the process of sustained technological innovation is localised are the following. First, the nature of knowledge, skills and collective routines

---

[179] (Storper 1997) p.19

[180] (Nelson and Winter 1982)

[181] One of the major insights of evolutionary economics (Nelson and Winter 1982) is that organisational routines represent the analogy of individual skills. Organisational routines are collective and complex patterns of quasi-automatic reactions driven by a set of pre-selected stimuli, operating on the basis of collective experiential learning, largely unconscious to the individuals involved. As skills, routines are acquired via collective learning by doing. The retention of routines is based on collective remembering by doing. In both cases, the *doing* does not require full consciousness, either at the individual (Squire and Kandel 1999) or at the collective level. The execution of a routine requires the spontaneous coordination and sequencing of a set of responses with a set of stimuli (which may come in any order). The communication system is based on a tacit language, full of locally understood words. Finally, the environment, in which the execution of the routine is carried out, is not separated by the routine itself, but constitutes the context in which the interpretation of the stimuli takes place. It is this contextual element of the organisational routine that creates the impossibility of reducing the routine to the sum of its individual agents' actions. Again, as for skills, the memory of the routines is stored in a distributed social network. The contextual dimension of organisational routines (as with similar organisational processes, such as communities of practice (Brown and Duguid 1991; Wenger and Snyder 2000)) makes them very difficult to manage, due to the inherent idiosyncrasy and diversity with which the elements of the routine become manifest.

is inherently localised. Second, because the package of competencies necessary for any innovation is not definable a priori, its concentration creates higher chances for the right package to emerge in an evolutionary manner. Consequently, the interaction between knowledge, skills, routines, cultural and organisational values tends to feed onto itself when geographically concentrated. Thirdly, the stickiness of technology diffusion creates a first mover advantage, which can be exploited to generate further innovation.

In this section I have suggested that the understanding of industrial clusters requires a complex set of disciplines, some of which are not directly related to the economics and sociology of industrial agglomerations. In particular, we need complexity theory and a dynamic theory regarding the morphogenetic properties of diversity. These are the objects of the next sections.

2.2.8 THE NEED FOR COMPLEXITY

The flexible specialisation school imposed to the attention of the world the industrial cluster as a novel form to organise production and innovation, by showing that economic analysis on its own does not manage to make sense of industrial clusters. The discovery of clusters places the spatial dimension at the centre of economic analysis. Territory brings with it history and local cultures, aspects of reality that neoclassical economics had considered as idiosyncrancies to be averaged out, rather than necessary elements of the phenomenon. But industrial clusters were only the tip of the iceberg. The real problem is to explain why economic activities with a high knowledge and creativity content are subjected to localisation. The first element of the puzzle is the complementarity (rather than mutual exclusivity) between flexibility and specialisation in distributed systems. This is paradoxical and constitutes the first indication that social and economic dynamics are very different in distributed systems. The second paradox is revealed by the transaction cost school by demonstrating that there is an alternative to minimising transaction costs by extension of the firm's control around high specificity assets. The alternative is unbundling of assets, disintegration of control and geographic agglomeration. But economics, sociology and geography do not provide all the answers. With the advent of the knowledge society the role of knowledge and technical change in shaping the structure of organisations has become paramount. The discovery that the dynamic of knowledge creation is strongly non-linear and subject to increasing returns leads to the recognition that innovation exhibits a spatial dimension. This aspect provides a rational for clusters and further input into the more general discussion regarding why agglomerations exists in the first place.

Although the frameworks described above have given outstanding contributions to the understanding of industrial clusters, still many questions remain open. Little attention has been devoted as to how the cluster as a meta-organisation achieves coordination and coherence. Or in other words, how can the cluster avoid degenerating into chaos in absence of a centralised controller? The coherence of the cluster is achieved by a sort of magic, whereby the parallel decision-making of independent multiple agents converge around some dynamic patterns. The problem to explain concerns the emergence of these patterns. The question can be formulated in this way. How does a collection of selfish agents, who base their decision-making on local information, and measure the effectiveness of the decision-making on the basis of short time scale considerations, build a complex system around the stratification of specialised and highly complementary knowledge? And, how does a collection of selfish agents prevent the dilution of the accumulated pattern of specialisation in absence of an externally given mechanism of coordination? The frameworks mentioned before are not really concerned with the question of emerging coherence. This is why the interpretation of the coherence of distributed system requires a theory that deals explicitly with self-organisation and with the conversion of '*micro-motives*' into '*macro-behaviours*'. This theory, complexity theory, is the object of the next section.

2.3 COMPLEXITY THEORY

"Einmal ist keinmal"[182]

2.3.1 THE ORIGINS OF THE IDEA

Complexity theory is concerned with the basic idea that the organisation of systems can be explained by means of emergent patterns of flows and interactions. From this point of view complexity represents an alternative system of explanation compared to

---

[182] "One time is no time" (my translation) (Kundera 1984). What is the role of experience if situations are never repeated?

epistemologies that explains the behaviour of a system by the analysis of its constituent parts. The contrast between these two systems of thinking can be traced back in history since ancient Greece. Parmenides and Hieraclitus were the first to represent the antithesis between *being* and *becoming*[183]. Hieraclitus was the first to stress the importance of context in the making of phenomena. "*On those who step in the same river, different and different waters flow*". If nobody can step twice into the water of the same river, (or as Kundera puts it: *"Einmal ist keinmal"*) because the flow of the water generates a different context, then explanations that assume the existence of a truth external to the phenomenon under observation are untenable. This contextual or holistic point of view will be rejected by modern science, which, starting from Galileo[184], espoused the Platonic principle of truth independent from observer and phenomenon. Building on Galileo's distinction between measurable quantities (legitimate objects of scientific inquiry) and non-measurable qualitative properties (legitimate objects of inquiries for humanities), Descartes developed the mechanistic metaphor of the world, in which inanimate and animate entities alike (with the exception of the human mind) behave as machines, subjected to mathematical laws. The Cartesian world is therefore rigidly deterministic. The behaviour of bodies can be described by means of analytical thinking, that is, by deriving the behaviour of wholes from the sum of the behaviours of its elementary components.

This approach clearly negates the existence of emergent properties. These are properties that emerge at certain levels of aggregation of parts, but do not exist (and can not be explained by) at the level of parts. For instance, the liquidity of water is an emergent property of the interaction of water molecules across a certain range of temperature, but it is utterly meaningless at the level of isolated molecules.

Newton[185] elevated the Galilean and Cartesian program to a universal description of the universe and built the foundations of what was to become modern science. During the period of Enlightenment all major sciences modelled themselves onto the basis of physics, including economics. Kant was probably the first philosopher to show the limit

---

[183] Being and becoming is the title of a famous book by (Prigogine 1980). A description of Heraclitean and Parmenidean systems is in Prigogine

[184] see for a discussion of the impact of Galilean thinking on science (Prigogine and Stengers 1997)

[185] (Prigogine and Stengers 1997)

of the mechanistic metaphor especially with regard to its applicability to the biological world. In his *Critique of Judgement* Kant writes of "*the previously unknown*" causality[186], referring to the behaviour of biological entities, where the parts seem to self-organise around a teleological plan, which is not external to the parts themselves. The parts exist not only in relation to each other, but produce each other by means of their interactions. This type of causality anticipates by almost two hundred years the circular causality of systems theory and Eigen and Schuster's hypercycle[187]. The organistic movement in biology will develop Kant's intuitions and search for an explanation of life phenomena in terms of patterns, flows and interactions of parts, instead of reducing the search to the ultimate reductionist causes. An example of the latter is Pasteur's theory of germs, whereby pathologies in animals are to be explained solely in terms of external germs[188].

In the 20th century reductionism will be seriously demolished as the only valid explanation, not only in social sciences but also in natural sciences. For instance, the "reformation" of physics ended up with the formulation of quantum mechanics, which put the last nail into the coffin of determinism by a) introducing probabilism, b) showing that ultimately there are no parts at all, c) introducing paradox and ambiguity at the heart of epistemology by claiming that entities show contradictory (and mutually exclusive) properties (wave/particle duality) and d) limiting knowledgeability by the Heisenberg uncertainty principle[189]. Yet, anti-deterministic quantum physics was still largely reductionist. At the opposite end of the natural science spectrum sits biology. Biology is dominated by Darwinism and offers a system of explanation based on *ex-post* rationalisation of chaotic events[190] similar to a narrative. There are no laws in biology, which share the same universality, applicability and forecasting power of physical laws[191].

---

[186] Indeed Kant wrote that the mechanism by which biological organisms organisms develop has "*nothing analogous to any causality known to us*" (Juarrero 1999) p.47

[187] see (Kauffman 1995)

[188] see for instance [Duris, 1999 #567]

[189] Heisenberg stated in his famous uncertainty principle that there is a limit to knowledge: knowledge of the simultaneous values of conjugated variables (such as time and energy, or position and momentum) is limited by the Plank constant. See (Feynman, Leighton et al. 1963)

[190] (Bak 1997), (Prigogine 1984), (Gould 2000)

[191] The "*survival of the fittest*" is valid in probabilistic terms in some cases (Cavalli-Sforza 1995), in others is only an oxymoron (Gould 1997)

Systems thinking and cybernetics are the two major contributions leading the way to complexity theory. Systems thinking originated with the works of Bogdanov in Russia and von Bertalanffy in Austria. Bogdanov was probably the first to envisage a universal theory of systems called *Tektology*[192], applicable to all fields of knowledge and based on universal principles of organisation. These laws of organisations allowed Bogdanov to distinguish between three types of systems: type A, in which the system is more than the sum of the parts, type B, in which the whole and the sum are equal, and type C, in which the sum of the parts is worth more than the whole. The distinction that I will extensively use in this thesis between aggregates and systems[193] is clearly contained in Bogdanov's work. Von Bertalanffy formulated a general theory called *General System Theory* that anticipated some of the achievements of Ilya Prigogine. The core of von Bertalanffy's theory rotates around the observation that biological systems do not seem to respect the second principle of thermodynamics, but on the contrary they seem to follow a plan of increasing order and complexity. This observation led von Bertalanffy to postulate that living beings were open systems whose behaviour could not be described by classical thermodynamics, but required a new type of thermodynamics[194]. But, if biological systems are not subject to equilibrium thermodynamics, this implies that regulation is not imposed by equilibrium laws and initial conditions, but on the contrary, is an inherent property of the system. It follows that biological systems are self-regulating. The departure from mechanistic and reductionist approach is now radical.

Systems theory evolved into cybernetics[195]. Cybernetics enlarged systems thinking to the new fields of machine control, neurology and neural networks, artificial intelligence and organisational behaviour. The central contribution of cybernetics is the concept of feedback loop. This is a set of causal links, arranged in a circular way, in which an initial change in one element causes a perturbation in the following element, which in turn is propagated in the chain until the cycle of causes and effects closes onto itself. Because the focus of cybernetics was on causal linkages and not on physical connections, the feedback loop became on organisational principle, which provided a solid and

---

[192] see (Capra 1996) pp.43-46

[193] see (Juarrero 1999) pp.109-111

[194] Almost half a century later, Ilya Prigogine developed that thermodynamics

[195] see (Wiener 1948)

mathematically sound basis to the concept of self-regulation. Self-regulation or homeostasis is achieved by the closure of feedback effects, which corrects excursion from the value at which the feedback loop balances. Cybernetics distinguished between negative and positive feedback loop. The former was thought to lead to a dynamic balance of self-correcting actions, the latter to a runaway reaction that would result in the destruction of the system. Feedback loops represented the first empirical demonstration of a self-regulating system, based on circular causality. Cybernetics discovered that the principle of organisation of a system was not represented by a set of laws (external and independent from the system itself) but was on the contrary based on the architecture of internal processes that guaranteed homeostasis thanks to their closure. Identity as *becoming* not identity as *being*.

Another major contribution of cybernetics was Asbhy's *law of requisite variety*[196]. If what makes a system is the closure of processes in a negative feedback loop, then systems are, as Ashby wrote: "*open to energy but closed to information and control*"[197]. A perfect example of such a system is the human genome, in which the process of change is blind to the external environment. A change in the genotype will create a different internal organisational architecture of genes expression, which may or may not confer a better fitness to the phenotype. However the change will not be driven by the environmental stimuli experienced by the phenotype. If the process of adaptation is blind to the environment, then survival requires that systems be supplied with a stock of reaction strategies that equals the environmental variety of stimuli[198].

Cybernetics focused on the problem of how order is maintained in complex structures and found that homeostasis is a result of the closure of feedback loops and variety control. The central question regarding how order is generated in the first place was instead addressed during the 60s and the 70s by a different group of scholars.

2.3.2 DISSIPATIVE SYSTEMS

---

[196] (Ashby 1956). This is explored in more details in section 2.5.3.5

[197] (Ashby 1956) p.4

[198] the situation is more complex as the phenotype has its own space of environmental adaptability. Moreover the new field of developmental evolution (see for instance final chapter in [Gerhard, 1997 #563]) is suggesting that typical genotipical traits exhibit a level of environmental plasticity. This suggests that some form of feedback ties the phenotype to the genotype.

Nobel Prize winner Ilja Prigogine has given one the most complete and sophisticated theory of self-organisation. The central concept in Prigogine's theory is that of dissipative structures. These are systems, which maintain their structure of orders by dissipating imported energy. Bénard cells[199] are the most famous example of dissipative structures. If a fluid is enclosed between two slabs and a controlled temperature differential is maintained between them, thermal energy will flow through the fluid. If the amount of energy is relatively small, then the system will dissipate the heat by conduction. This is a system near equilibrium. If the amount of energy exchanged overtakes a critical value, then the system goes through a phase transition in which a completely new pattern of organisation emerges for the dissipation of energy. The spontaneous co-ordination of billions of molecules gives rise to a pattern of hexagonal convection cells. Prigogine made the startling discovery that taking the system away of equilibrium (by forcing it to exchange energy with the environment) gives rise to completely new and more sophisticated structures of order. The way in which new order arises reveals a major step forward from cybernetics. Cybernetics concentrated on negative feedback loop in order to explain homeostasis via self-regulative mechanisms. Models based on self-organisation instead discovered that positive feedback loops played a fundamental role in the order creation phase (when a system goes through a phase transition). In correspondence of a specific amount of energy exchange, some naturally occurring reactions are selectively amplified to the point of taking over the whole system. In the Bénard cells this mechanism is the upward or downward force that acts at the level of single droplets as a consequence of temperature induced density change. When the first droplet wins the viscosity resistance, it will stimulate others to behave accordingly, triggering a cascading reaction that will take over the whole system. Interestingly, this force acts only at the local (molecular or droplet) level. The correlation of billions of molecules that gives rise to the Bénard cells is an emergent property at the macro-level. "*Emergence occurs when interactions among objects at one level gives rise to different types of objects at another level. More precisely, a phenomenon is emergent if it requires new categories to describe it, which are not required to describe the behaviour of the underlying components*"[200]. The new organisation of the system is therefore

---

[199] (Nicolis and Prigogine 1989)

[200] (Gilbert and Troitzsch 1999) p.10

based on some dynamic processes that were already present in the system, but did not play (before the phase transition) any major role[201]. The transition from conduction to convection is sudden and involves a radical restructuring of processes and structures at the macro level. A phase transition is therefore radically non-linear and threshold dependent.

A phase transition that leads toward an increase in complexity is accompanied by a bifurcation in which several alternative stable configurations are available to the system. In the Bénard cells only chaotic noise will determine whether the sense of rotation of the cell will be clockwise or counterclockwise. This means that the system can exist in two possible states, one in which the order of the cells starts with a clockwise cell and an alternative one in which the starting cell is counterclockwise. The selection of any of them depends on local chaotic conditions or instabilities that are present in the system. Importantly, noise and micro-instabilities are random events that resist any attempt of predictability. Although this point may seem insignificant, it represents a defeat for determinism[202]. The appearance of bifurcations in correspondence of a phase transition bears two important consequences. First, as Luhman would put it: "*complexity entails that, in a system, there are more possibilities than can be actualised*"[203]. The trajectory that a system follows is only one of the possible ones. This is a final blow to determinism as it introduces an evolutionary dynamics and with it strong process irreversibilities in the nature of physical processes. Second, making even physical processes more similar to biological ones, Prigogine introduces a *narrative* element in the description of nature. Re-conducing all sciences to a form of history allows Prigogine to claim that complexity goes beyond the traditional separation between natural and social sciences and the alienation[204] that the Galilean program had created. The conflict between what J. P.

---

[201] In the words of Nicolis and Prigogine: "*non-equilibrium reveals the potentialities hidden in the nonlinearities, potentialities that remain dormant at or near equilibrium*" (Nicolis and Prigogine 1989) p. 60

[202] (Nicolis and Prigogine 1989) p.14

[203] In (Cilliers 1998) p.2

[204] (Monod 1972). Also Laing wrote: "*Galileo's program offered us a dead world: out go sight, sound, taste, touch, and smell, and along with them have since gone aesthetic sensibility, values, quality, soul, consciousness, spirit. Experience as such is cast out of the realm of scientific discourse. Hardly anything has changed our world more during the past four hundred years that Galileo's audacious program. We had to destroy the world in theory before we could destroy it in practice*" R.D. Laing, quoted in (Capra 1988) p.133

Snow[205] described as *the two cultures* (the ones associated with social and natural sciences) can be overcome thanks to the freedom introduced by the creation of order in complex systems. The central intuition by Prigogine, that non-equilibrium situations can be a source of emergent order and that the description of non-equilibrium requires non-linear mathematics, opened a completely new avenue of research that has culminated in complexity theory.

2.3.3 THE HYPERCYCLE

Prigogine showed that the morphogenetic properties of dissipative systems depend on the capability of amplifying fluctuations (that at equilibrium are irrelevant) or instabilities to the point at which a new form of order emerges. Around the same years biochemist Manfred Eigen[206] (Nobel Prize for chemistry) was working on the problem of the origin of life. Eigen postulated that the emergence of life from non-living systems could be explained by the concept of the hypercycle. This is a set of catalysed[207] chemical reactions, which achieves closure. If we imagine that chemical component *a* can catalyse the production of component *b* and *b* catalyses *c* and so on, a finite probability exists that (down the chain) component *j* will catalyse the production of *a* itself. When this happens the cycle becomes a hypercycle characterised by some fundamentally new qualities:

1) The hypercycle is a giant self-sustaining and self-organising chemical feedback loop

2) The hypercycle represents a qualitatively different system from the chemical "soup" from which it emerges. The closure of the set creates a self-sustaining mechanism that at the same time a) establishes a boundary around the newly emerged system, b) provides an organising principle for the internal dynamics of the set, c) fixes rules by which certain components are amplified (the catalysed reactions) and others are selected out, and finally e) filters the information/components allowed in the cycle and rejects the others.

---

205 (Snow 1959)

206 (Eigen 1971)

207 A catalyst accelerates a chemical reaction by reducing the activation energy barrier (Nicolis and Prigogine 1989)

3) The new system is not a physical system in the traditional sense of the word. It is informationally closed but energy and matter open. The organisational rules that drive the formation and the overall dynamics, the rules that keep the system in balance and make the system able to withstand perturbations, are all internal to the system. They emerge with the system and in a way they are the system.

4) The closure of the set of catalytic reactions constitutes a phase transition that discriminates between two fundamentally different forms of existence of the chemical "soup". In fact, although the single reactions act to take the system away from equilibrium, it is the closure of the set that acts as a collective catalyst to the chemical "soup", and, for the reasons mentioned in points 2 and 3, it is the closure that transforms the "soup" in a different system (or even better by providing an organisational principle turns the "soup" from an aggregate into a system[208]). Interestingly the closure of the system is not local, but is a collective property of the hypercycle[209].

Newtonian science based on the Aristotelian efficient cause is at odd to describe such a system. Why? For two main reasons. First, efficient causality requires that the cause be external to the effect, that is, exist independently from the object of causation. This is true for the single steps of the catalytic cycle but patently untrue for the hypercycle itself, for which we have to accept a type of circular causality[210]. Second, Newtonian science is reductionist, or in another terms based on analytical processes whose purpose is the identification of the first components and laws from which the machinery of efficient causality start. As Bak and Chen[211] point out:

" *traditionally, investigators have analysed large interactive systems in the same way as they have small, orderly systems, mainly because the methods developed for simple systems have proved so successful. They believed they could predict the behaviour of a large interactive system by studying its elements separately and by analysing its microscopic mechanisms individually. For lack of a better theory they assumed that the response of a large interactive system was proportional to the disturbance*".

---

[208] see next section for a discussion of aggregates and systems.

[209] (Kauffman 2000) p.105

[210] (Juarrero 1999) points out that the understanding of complex systems requires all four Aristotelian causes. Moreover Kant himself in his Critique of Judgement recognised that biological entities escape the logic of efficient causalism and necessitate a "*previously unknown*" form of causation. This "*previously unknown*" cause has to take into account the self-organising properties of life.

[211] (Bak 1997) p.21

This approach does not work with the hypercycle. The mix between bottom-up (each reaction is defined locally as it is only connected to the previous and following one) and top-down properties (the closure is a non local feature, which determines the relative abundance and distribution of the various molecules, and more in general, the organisational principle of the system formed around the hypercycle), escapes the logic of a reductionist approach, which misses completely the emergent properties and the local non-local mixof the hypercyle.

Summing up, complex systems originate through a phase transition in which the organisation of the system changes dramatically and a new type of attractor[212] emerges. The new type of order, more elaborated and complex than the previous one, involves some type of correlation between distant and autonomous parts. This correlation is provided by the closure of micro-mechanisms of interactions, such as the hypercycle.

2.3.4 AGGREGATE, SYSTEM, STRUCTURE AND ORGANISATION

I intend in this section to discuss a major difference between complex and non-complex systems, a difference that will play a major role in this thesis. We have seen that the dynamic of complex systems shows an endogenous nature, i.e. it emerges around the selective amplification of some spontaneously occurring set of micro-phenomena around which new structures of order crystallise (often in the form of a hypercycle). The boundary and organisation of the new system are purely dynamic and defined by (and consists of) a set of processes, sustained by the non-equilibrium conditions. This endogenous organisation is unique (because of the role of instabilities and random events in the phase transition) and allows us to draw a fundamental distinction between aggregates and systems. The former consists of a set of parts whose properties do not depend on the specific context in which they happen to be. The latter instead consists of parts organised into a set of multiple dependencies and feedback loops in which there is a bi-univocal relationship (bottom-up from parts to whole and top-down from whole to parts) between the parts and whole. Rescher[213] gives the following definition: "*the root idea of system is of integration into an orderly whole that functions as an organic*

---

[212] for the concept of attractor and dynamic stability within attractor's range see (Allen 1997)

[213] (Rescher 1979) p.4

unity". An example of whole to parts dynamic is the hypercycle's collective organisation (provided by the closure of the catalytic reactions). This exerts a top-down selection on the concentration and diversity of molecular species, thereby enabling certain reaction channels and excluding others. As concentration and availability of molecular species can give rise to new molecules and new mechanisms of interactions, clearly the systemic effects of the hypercycle affects the properties of the components of the hypercycle itself. This type of influence of the context on the parts is an example of a system in action.

The fact that a system can be described as a functional unit implies, according to Maturana and Varela[214], that any system has an organisation and a structure. The former consists of "*the relations among components that constitute a composite unity as a unity of a particular kind*"[215]. A structure consists instead of "*the actual components and the actual relations among them that at any instance realise a particular composite unity as a concrete static or dynamic reality*"[216]. A dead car has still a structure but no organisation. This distinction can be extended to internal and external structure. The former refers to the parts and relations between parts that belong to the entity under consideration. The latter refers to the set of interactions between the entity's parts and the environment. This distinction is important because "*the environment in question to the system's theorist is not the total environment but the environment that affects and is affected by the thing in question*"[217]. The concept of environment changes from the static Darwinian environment, which acted as the background for the evolutionary struggle, to the set of things in dynamic coupling with the structure of our entity. The environment becomes then a relational property of the entity under consideration, linked with it through the entity's external structure. The effects of the system-environment coupling becomes manifest in the process of coevolution, which expresses the parallel reciprocal transformations between the entity and its environment through the multiple channels of interactions that constitute the system's external structure.

---

[214] (Maturana 1998)

[215] Cited in (Juarrero 1999) p.109

[216] Cited in (Juarrero 1999) p.110

[217] (Juarrero 1999) p.110

The distinction between organisation and structure allowed Maturana and Varela to explicitly formulate the difference between allopoietic and autopoietic systems. In the former the organisation is supplied from the outside, that is, it requires the presence of an external designer. Machines are allopoietic. Living systems instead are autopoietic, literally, *self-making*. What Maturana and Varela add to the theory of complex systems is their idea regarding the role of circular causal processes in networks. The organisation of networks revolves around a set of reciprocally feeding causal processes whose first purpose is to maintain itself. Maturana would claim that the organisation of systems *"allows for evolutionary change in the way the circularity is maintained, but not for the loss of the circularity itself"*[218].

This step represents a novel approach to the study of systems. In fact if we accept the principle that many systems are autopoietic, then the problem of how these systems interact with their external environment becomes real. Maturana and Varela's answer is that the process of interaction is really a process of cognition. The autopoietic system modifies its patterns of circular processes in full autonomy from the environment, in much the same way as a neural circuit generates a response to an external stimulus by changing the patterns of connections and weights, without building any local representation of the external object. What Maturana and Varela suggest is that self-organisation implies self-referral in the way in which cognitive processes take place[219].

---

[218] cited in (Capra 1996) p.96

[219] As neuro-scientists Maturana and Varela were both interested in the problem of perception and managed to link complexity and perception into a unified whole. Their autopoiesis principle of organisation implied that the mechanism of perception is not to be seen as a "*representation of an external reality*", that exists independently from the observer, and is (in some neural forms) reproduced within the observer. On the contrary the mechanism of perception is nothing else than a transformation of patterns in the circular processes, an alteration in the mechanisms of organisation, which may or may not improve the fitness of the organism. This is not to say that the system and the environment do not interact, but is analogous to the distinction between genotype and phenotype. The phenotype represents the expression of the genetic information contained in the genotype in a context mediated by environmental constraints. However, environmental stimuli have no effects on the genotype (literally blind to the environment), which is instead driven by random mutations. The organisation of the genotype and the changes that take place within are self-referring and autonomous. The gap between environment's actions and evolution of the organisation of the living leads to the fact that systems have to refer to themselves in the process of evolution. We find then that living systems are not only self-organising but self-referring. In a way living systems specify their external environment, instead of representing one, by means of changes in the patterns of circular causality. This process, that Maturana and Varela call cognition, coincide with the set of processes that they define as the organisation of the system. Life itself is therefore a process of cognition and self-organisation by self-referral.

2.3.5 CONSTRAINTS AND MORPHOGENESIS

In the previous parts of this section on complex systems I have shown using the relevant literature that a) the formation of a complex system is endogenous, b) a complex system has an organisation and a structure, c) a complex systems' environment is specific to the system d) the environment is in a coevolutionary coupling with the system itself, and finally e) living systems are special types of complex systems based on the process of cognition. What needs clarification is the dynamics of formation of complex systems. Following Juarrero[220] I will introduce the concept of constraints as generators of order (morphogenesis). The traditional interpretation of constraints, coming from classical science, is that of a limit to motion and/or reduction in the number of degrees of freedom. A pendulum is constrained along a certain trajectory, with the constraint acting to force the motion along a specific type of trajectory. A constraint can be seen as an environmental limit to the freedom of an object. But in the light of the previous discussion on complex system's genesis, constraints can be seen as the manifestation of a connection between different objects, thereby providing the first step for the transition from an aggregate to a system. *"Constraints are therefore relational properties that parts acquire in virtue of being unified – not just aggregated – into a systematic whole"*[221]. The pendulum becomes a new object whose existence and dynamics is due to the action of some mechanical constraints. The rod that connects the fixed point to the mass becomes part of the system and, by virtue of restricting the freedom of the mass, generates an object that can, for instance, measure time. Constraints have therefore a generative effect. Seen from this point of view, constraints act effectively as linkages between parts, on the one hand constraining the possibilities of action of the single parts and, on the other hand, defining an architecture of interactions which effectively fixes the structure of a network[222] into a specific configuration. It is this architecture that allows the system to perform new and more sophisticated functions.

---

[220] (Juarrero 1999)

[221] (Juarrero 1999) p.133

[222] (Campbell 1982) p.134

Constraints come in two different forms: context independent and context sensitive. Context-free constraints[223] act in the same way on all elements of a population. For instance, the probability distribution of letters depends on the specific language. For instance, the frequency of certain groups of letters, for instance *tion*, is high in English but low in Italian. The constraint here acts to remove the group *tion* from the situation in which any group of four letters is equally probable (by making *tion* more probably than, say, *zion)*. This type of constraint has no generative property insofar it exert a selectionist pressure on the probability of occurrence of an event without generating new forms of connections. A context sensitive constraint (CSC) instead generates a conditional probability for the occurrence of an event, which depends on the occurrence of a previous event. In this way context dependent constraints link successive events to a specific context. Because the order in which the events take place is temporally dependent, then context sensitive constraints introduce the dimension of time and memory in the system. "*Once the probability that something will happen depends on that is altered by the presence of something else, the two have become systematically and therefore internally related*"[224]. Context sensitive constraints also generate higher level grammars. By linking single letters in *i*-tuplets (the group *qu* is an *i*-tuplet*)* CSCs create a much larger space of possibilities for creating new entities by mixing *i*-tuplets together. As in a Lego construction kit, the combinatorial space and the complexity of the constructions increase exponentially with the number of elements to be combined and with the dimension of the single blocks. The action of CSCs come in two different streams, bottom-up and top-down. Bottom-up CSCs link simple events, for instance by increasing the probability of the letter *u* after a *q* or favouring an increase in concentration of a chemical component thanks to the presence of a catalyst. The second type (or top-down constraints) acts instead to select the type of environment in which the single events take place. An example of top-down constraint is the hypercycle. The collective organisation of the hypercycle creates a range of boundary conditions with respect to, for example, which molecular species are allowed in the cycle and which are selected out. Because top-down constraints are not locally definable, they act as systemic properties of the whole. Top-down constraints are effectively emergent properties

---

223 The definition is by Gatlin. Cited in (Juarrero 1999)

224 (Juarrero 1999) p.139

caused by the interactions at the lower level. The fascinating aspect of this approach to complexity (often missed in the so called *simple rules* approach[225]) is the interaction between the bottom-up direction, generating the system, and the effect of the system on the components of the system itself. Once the organisation of the system is generated, this takes over and constraints the behaviour of the component parts.

A related aspect of constraints is developed by Stuart Kauffman in Investigations[226]. Kauffman points out that traditional science hides the generative roles of constraints in the formulation of the initial conditions of any problem. Let me give an example. A situation of non-equilibrium will, by default, be characterised by an anisotropic and non homogeneous concentration of energy (or information). Any simple device that performs work exploits some local difference in energy between the parts of the system or the system and its environment. What science gives for granted (or hides under the term initial conditions) are really two things. First, the capability to recognise the presence of local non-equilibrium conditions requires the presence of an apparatus, which is at least as complex and sophisticated as the source of the non-equilibrium to be exploited. Second, the fact that, in order to perform work (in Atkins' words "*the constrained release of energy*"[227]), a system must have spent energy to build those constraints that are able to a) spot the dishomogeneity in the distribution of energy and b) exploit that source of non-equilibrium. We are back to circular causality. To extract energy, constraints must be in place, but to build those constraints energy has to be spent. Kauffman's approach to work, energy and constraints casts light on the morphogenetic properties advocated by Juarrero in her discussion of constraints. The relational property of constraints, achieved by linking different parts of an aggregate into context-sensitive stable system of relationships, constitutes the building up of the channels along which energy can be more efficiently extracted to perform work. The rather mysterious aspect of order generation, implicit in Juarrero and systems theorists, finds here a '*mechanistic*'[228] interpretation. Relating aggregate's parts into systems allows the utilisation of energy into work, energy that was previously disposed as entropic waste.

---

[225] (Resnick 1997). For a critical discussion of the simple rules approach see (Cilliers 2000)

[226] (Kauffman 2000) pp.96-98

[227] Atkins cited in (Kauffman 2000) p.97

[228] I use mechanistic as Maturana and Varela do (Maturana 1998), meaning that it does not require the use of metaphysical or *elan vital* type or explanation

Furthermore the correlation permits to use the newly performed work to build more constraints to refine the conversion of energy into work. The system becomes self-generating and self-sustaining. A further implication regarding the relationship between increasing diversity and generation of constraints will be examined in the next section.

This section has described the emergence of a more general view than reductionism. According to complexity thinking, reductionism is a method of inquiry in the analysis of systems, valid when, as Bogdanov[229] claimed, the whole equals the sum of the parts. A reductionist analysis is usually appropriate for systems at (or near) equilibrium that are usually characterised by linearity, absence of feedback loops (streamlined connectivity) and lack of self-organisation. These systems have been the objects of scientific and philosophical analysis for the past few centuries. The astonishing successes of modern science and engineering are there to witness to the potency of reductionist analysis and to the amount of knowledge gained in the description of such systems. Much less is known about the behaviour of non-linear systems. The sections that follow will describe some general properties of self-organising systems.

2.4 POWER LAW PHENOMENA

2.4.1 THE RANK-SIZE RULE

In the 30s Zipf[230], a Harvard linguistic professor, discovered a curious phenomenon[231]. Plotting (on a double logarithm graph) the rank of the American metropolitan cities against their population, the graph showed a straight line with slope of almost exactly -1[232]. With time it came to be realized that the Zipf example was far from being isolated. The rather large class of phenomena[233] that on a double logarithm paper approximate a straight line are known as obeying a power law.

---

[229] see section 2.3

[230] (Adamic)

[231] (Krugman 1997)

[232] It basically means that for each city of rank 1 with a population a million, the second ranked has a population of 500000, the third ranked one third of the first etc.

[233] In physics the class of phenomena that obey a power law are also known as 1/f (Bak and Chen)

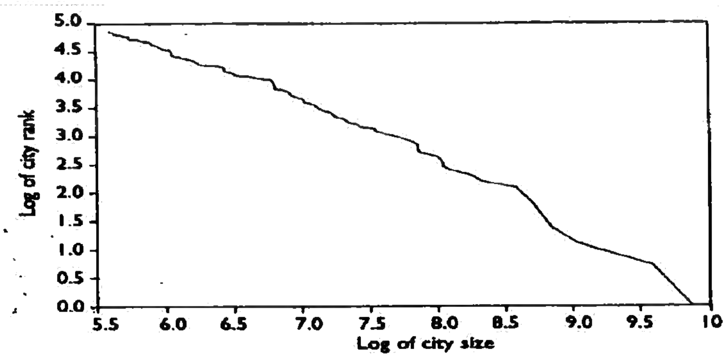


Figure 3.1 Zipf's Law: The size distribution of metropolitan areas in the United States is startlingly well described by a power law, with an exponent very close to 1.
Source: *Statistical Abstract of the United States*, 1993.

Figure 17 - The Zipf law (from Krugman (1996) p.40)

This regularity is persistent in time and space. The same set of cities is power law-consistent from 1860 to 1991[234]. The same behaviour appears in other countries, as well[235]. Explaining this regularity has been a challenge for economists and geographers for a long time. As Krugman puts it: "*we are unused to seeing regularities this exact in economics – it is so exact that I find it spooky*"[236]

Two explanations have been proposed: hierarchical central place and dynamic.

The hierarchical explanation goes back to Christaller and Losch lattice theory[237]. According to this theory, urban centres show a hierarchical distribution of city sizes due their different specialisation patterns. At one end of the spectrum there are small centres, specialising in farming and supplying larger centres. At the other end we have large cities, specialising in industrial activities and administration, and completely dependent upon smaller cities for agricultural supplies. Alternatively, we can imagine a managerial hierarchical structure, whereby for each central office there are (for instance) three subsidiaries, which in turn supervise other 3 subsidiaries, and so on. These approaches are consistent with a power law but are too simplistic. For instance, if the distribution of city sizes were dependent on a supposedly higher dependence of smaller centres on farming, then, the regularity should have disappeared with the marginalisation of

[234] (Krugman 1997) p.256

[235] (Krugman 1996) p.40

[236] (Krugman 1996) p.40

[237] (Fujita, Venables et al. 1999); (Allen 1997) p.28-30

agriculture. Agricultural contribution to GDP has dramatically declined over the last 50 years[238] and therefore it is not clear why the distribution of city/town/villages has not been affected. In the case of managerial structures there is simply no regularity in the distribution of managerial posts. The constancy of relationships between levels does not hold true[239].

Alternatively the rank-size rule might emerge as a result of a random dynamic behaviour of growth. Three conditions have to be in place:

1) the distribution under study is subject to growth (the growth process can be positive or negative);

2) growth is spatially random, that is, there is no preference for any part of the agglomerate;

3) growth depends linearly on the size of the node.

If these conditions hold, then it can be demonstrated[240] that the phenomenon obeys a power law.

What type of growth are we likely to obtain? First, growth obeys a fractal distribution[241]: the size-independence growth rule makes the same type of pattern appears at different scales, which implies that different parts at different level of aggregation are self-similar. This is the signature of fractality. Second, the same rule excludes any mechanisms of growth based on increasing returns. It follows that the parts will grow at a rate proportional to their size.

An interesting model based on the above rules was proposed by Nobel laureate Herbert Simon (1955). Let us start from a lattice. Each point may represent a town, a firm or anything discrete. Let us also imagine that some points be occupied by towns of a certain size, others instead be empty. We now impose a dynamic condition: a '*lump*' (for instance, a certain number of people) has a probability ϕ of coalescing to an existing town and a probability $1-\phi$ of starting a new town. Given these simple rules (which

---

[238] contributing nowadays around 6-8% in industrialised countries (Cipolla 1970)

[239] (Krugman 1996) p.42

[240] (Krugman 1996) pp.92-4

[241] (Casti 1994) (Gleick 1987)

respect the three conditioned described above), the resulting distribution is indeed a straight line on a log-log graph.

In opposition to the hierarchical central place model Simon's model makes no assumption and does not depend on any relationship between different economic sectors.

The question is: can such a simple mechanism describe surely complicated aggregation and growth phenomena? To answer this question, we need first to describe a complementary class of models, which go under the name of self-organised criticality models.

2.4.2 SELF-ORGANISED CRITICALITY

This class of models is symbolised by another idealised (but feasible) experiment. The experimental apparatus is very simple: a sand-pile. If we start dropping sand, this may form a new pile or add to an existing one. With time the heaps will coalesce until a sand-pile is formed. At this point something interesting happens: the pile will reach a critical slope and self-organise its dynamics around that slope. Each new grain will either rest somewhere on the pile or, by rolling down the slope, start an avalanche (at minimum of size one). Measuring the size of avalanches (for instance its mass) and plotting the results against the frequency of avalanches of the same dimensional class gives again a straight line on a double log graph. A way to explain the avalanche is the following: each new grain has a probability ϕ to come to rest and $1-\phi$ to displace an adjacent grain. If the latter happens, then the two grains have a probability $(1-\phi)^2$ to displace other two grains. Because the phenomenon is random, the probability per grain of sand does not change. The probability of getting a bigger avalanche simply scales with the number of particles involved. Mathematically this means that the behaviour of the avalanches obeys a power law of the type:

$$F = \beta * S^{\alpha}$$

Where F stands for the frequency of avalanches with size S and β is a constant.

At the critical state the system tunes itself towards a critical slope, whereby any external perturbation (the dropping of an additional grain of sand) induces a system's reaction that can span any order of magnitude with a frequency distribution expressed by a power law. This behaviour is counter-intuitive as we generally assume a linear relationship between the size of the perturbation and the size of the effect, i.e. small causes yield small effects. Indeed this is true before the self-critical state is attained. At the self-critical state instead the proverbial straw can break the camel's back! A first implication has to do with catastrophe theory. For instance, the conventional explanation regarding the extinction of dinosaurs at the end of the Cretaceous is imputed to a giant asteroid and/or massive eruptions[242]. Instead, according to a self-critical model, internal causes may have been in action, causes that were progressively amplified until a catastrophic chain reaction took place[243]. Fundamentally self-organised criticality (SOC) is a theory about endogenous sources of change in a system.

This point leads to a fundamental difference between complexity-based theories and other theories, that is, the issue of holism and reductionism[244]. The fact that a self-critical systems spontaneously tunes itself towards a self-critical state[245], where "*the system organises itself towards the critical point where single events have the widest possible range of effects*"[246], makes the reductionist approach inappropriate for the study of self-critical systems. Fundamentally self-critical behaviour takes place when an internal system of linkages makes the behaviour of an agent dependent upon the behaviour of others, so that small and local fluctuations can be amplified according to a power law distribution to achieve systemic effects. Due to the emergent properties of the interconnected whole, the study of isolated elements or interactions will cast only a partial light upon the system's behaviour unless a holistic approach to the study of complex system takes place in parallel.

Sand-pile mechanisms have turned out to be very common in nature[247]. From the dynamics of earthquakes[248] to the paleontological records of species extinction[249], from

---

[242] (Raup 1999)

[243] (Gould 2000), (Raup 1999) pp.217-218

[244] see citation by Bak and Chen at p.28

[245] (Bak and Chen 1991; Kauffman 1995)

[246] (Cilliers 1998) p.97)

[247] A recent book of popular science, Ubiquity (Buchanan 2000), is entirely devoted to the pervasiveness of self-organised critical phenomena

[248] (Waldrop 1992) pp.305-306

[249] (Raup 1993; Bak 1997; Raup 1999)

the succession of booms and busts in economic cycles[250] to the dynamics of artificial life games[251], there seems to be a common dynamic across disparate fields that regulate the relationship between frequency and size of events in a self-organised critical manner.

2.4.2.1 The Scheinkman-Woodford model

An interesting application of SOC to economics is the Scheinkman-Woodford (1995) model (SWM)[252]. The starting question has to do with the observed correlation between fluctuations in input factors of large economic systems (such as consumers buying patterns or production techniques) and the overall reaction of the aggregated system. Conventional logic, based on the concept of homeostasis or equilibrium, would foresee an averaging out of small fluctuations in the large aggregated system instead of a correlated macro-reaction.

The SWM is very simple: it represents an idealised economy where N layers of M firms supply and buy from each other. At the top of the matrix sit the final assembler firms, at the bottom the providers of raw components. The rules of engagement are elementary. Each firm buys from the lower layer and sell the to upper layer. When a unit order arrives[253] from a customer (this represents the exogenous shock or fluctuation), the firm checks its inventory. If the good is available, it is sold to the customer in the immediate upper layer. If not, two orders are passed down to two neighbours of the supply chain, enough to produce two goods, of which one is sold and the other stocked. This rule corresponds to a real cost in holding stock and leads therefore to minimisation of inventory. Initial distribution of stock is randomly generated.

---

[250] (Krugman 1996)

[251] (Buchanan 2000)

[252] (Scheinkman and Woodford 1994)

[253] Indivisibility in production techniques can be used to justify the assumption of unit order

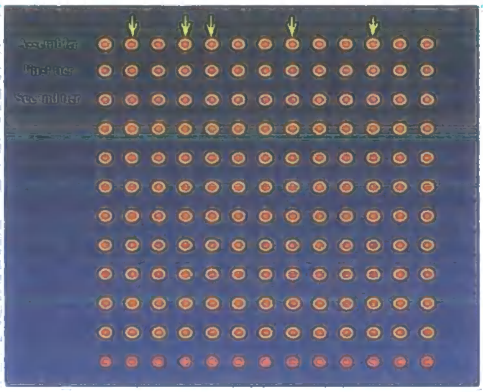


Figure 18 - 1st case: high level of inventory

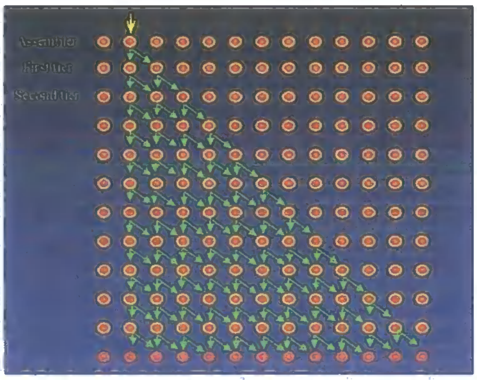


Figure 19 – 2nd case: low level of inventory

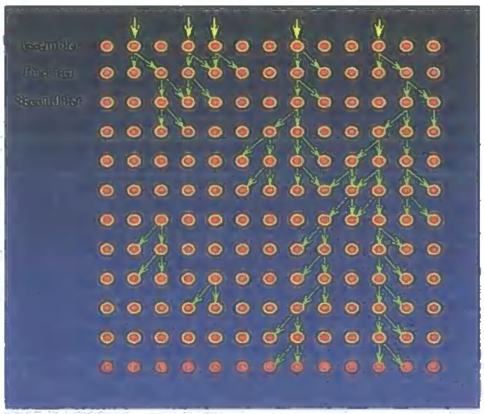


Figure 20 – 3rd case: random level of inventory

What happens is very simple. When an order arrives, if the firm involved has a high level of inventory, the order is satisfied by the front line and the fluctuation is rapidly absorbed. If instead the level of inventory is low, the initial order triggers two orders down the chain, which causes four orders in the third line and so on. The initial unit order (fluctuation) gets amplified until it becomes a cascade with the potential to involve the whole system in a long series of transactions. However, the simulation shows that after a transient, the system settles down to a self-critical level corresponding to a power law distribution of transaction lengths. The system tunes its inventory level spontaneously to the self-critical situation. Short chains are very frequent, long chains are less frequent with the probability of chain length being controlled by a power law formula. Importantly long chains (large effects) do not require correspondingly large fluctuations. As in the case of the sand-pile the overall state of the system can be described in terms of a single parameter, which in the SWM is the average level of inventory, whereas in the sand-pile it is the slope. This is remarkable, as there no are no specific macro reasons that push the system to react with a power law distribution to incoming fluctuations.

In the following section I will show the importance of the SWM model in linking SOC behaviour to the *edge of chaos* dynamic.

2.4.2.2 SWM and edge of chaos

The dynamics of our idealised economy depends on the parameter inventory level. There are broadly three dynamics set by parameter:

1. High inventory parameter: the system is static; fluctuations are rapidly absorbed by satisfying the orders; chains are very short (Figure 18).
2. Low inventory parameter: the system becomes chaotic; very few orders are satisfied by the first layers and the chains become potentially indefinite (Figure 19);
3. Inventory parameter tuned: the system sets in a situation intermediate between static and chaotic, where chains of any length appear, with islands of transactions propagating in the supply chains. This is the renowned "*edge of chaos*"[254] situation typical of complex systems (Figure 20).

What the SWM tells us is that SOC is the distinguished feature of systems on the edge of chaos. This is a major step forward. In the Bénard cell experiment, the forming of convective cells is dependent upon the tuning of an external parameter, the temperature differential. That is, the state of the system is controllable from the outside. The

---

[254] (Waldrop 1992; Goodwin 1994; Kauffman 1995)

emergence of correlated macro-scale behaviour is only partially emergent because of the constraint exerted by the external control bias. In the SWM instead, the web of connections and the non-linearities of the interactions spontaneously bring the system into a self-critical behaviour that, almost by magic and without outside forces, reveals an *edge of chaos* situation. The immediate consequence is that we can use the SOC as the footprint for identifying the dynamic state of a system under the conditions (internal and external, structural and environmental) in which it is operating. Also, if the three conditions identified in the discussion of the *rank-size* rule apply, then we have a way to characterise the type of interactions that leads a system on to *the edge of chaos*. The question then becomes: are the size-rank rule and self-organised criticality two manifestations of the same thing, or not?

Before trying to answer this question I need to introduce a further power law model: the *scalefree networks* model.

2.4.3 SCALE-FREE NETWORKS

The family of theories studying phenomena obeying a power law has recently added a new member: the scale-free network (SFN) theory[255]. Barabasi and colleagues started their investigation from the question whether the nodes of the Internet were characterised by a number of links distributed according to a random distribution[256]. In place of the hypothesised random distribution, they found instead a power law. This implies that there is no typical node, exhibiting a typical number of links and therefore determining the scale of the system. Instead what they found was that most nodes have few links and very few of them - the hubs - exhibit a disproportionate number of links,[257]. The work by Barabasi and others[258] has demonstrated that SFNs appear in fields as

---

[255] (Barabasi 2002)

[256] this is a class of networks studied by the Hungarian mathematician Paul Erdos (Barabasi 2002) that assumes that links are randomly distributed across nodes. The distribution of links per node is mapped by a bell-shaped distribution, whereby most nodes have a typical number of links with the frequency of remaining nodes (with number of links above and below the average) rapidly decreasing on either side of maximum.

[257] Let's, for instance, imagine we have a network with 1000 nodes obeying a power law of 2. The number of nodes with 2 links is simply the total number of nodes 1000 times the inverse of 2 to the second power, that is 250. The number of links having 100 nodes is $1000*(10^2)^{-1}$, that is 10.

[258] (Réka, Jeong et al. 1999; Jeong, B. Tombor et al. 2000), (Barabasi 2002)

disparate as epidemiology, structure of the Internet, metabolism of cells, genetic circuitry, and social networks.

This type of structure has a profound influence on the networks' behaviour. These networks are generally very stable and robust. For instance random networks are vulnerable to random attacks (that is attacks that remove nodes in a random way), whereas SFNs are not. As most nodes have very few links, any random attack has a high probability of removing uninfluential nodes without compromising the architecture of the network. But malicious attacks, writes Barabasi, can take the system down by destroying the hubs, in the case of the Internet, the Yahoos, CNNs and eBays type of nodes[259].

SFNs seem to self-organise when two basic conditions are present: first, the network is growing, and second, new links follow a preferential attachment rule. This *"rich get richer"*[260] effect indicates that links being added to the network attach preferentially to already rich nodes, with a probability depending on the existing number of links. These conditions are sufficient to determine a power law distribution[261].

The SFN analysis focuses on the topology of networks. It is, from this respect, very different from the rank-size rule and self-organised criticality models. Yet, the same type of distribution, the power law, pops up again. I will discuss the implications of these three classes of models in the chapter on theory.

2.4.4 SOME CONSIDERATIONS ON POWER LAW BEHAVIOUR

I have in this section rewieved three independently proposed models/evidence on the behaviour/structure of some types of networks. There are three further aspects about power laws that deserve attention.

First, the interesting question arises whether the similarities between the three models reflect some underlying pattern and whether ultimately they reflect a universal dynamic across self-organising systems. I will discuss this point further in chapter 3. Second, I

---

[259] This property can be used in epidemiology to stop the diffusion of epidemics by addressing the hubs (Barabasi 2002), Chapter 10

[260] (Barabasi 2002) pp.79-92

[261] Barabasi does not seem to be aware that his two 'laws' are a riformulation of Simon's conditions expressed in his 1955 paper on 'lumps and clumps". See my discussion in section 2.4.1

have tried to show, using evidence from complexity, economic geography, economics and physics, that there is a close link between self-organising systems (or systems on the *edge of chaos*) and power law behaviour. This suggests that power law can be used as a tagging device to identify the dynamic state of systems (that is, whether a system is self-organising). This point will be developed further in the chapters on theory and empirical results. Third, a system obeying a power law spans all orders of magnitude across the chosen variable. This fact seems to suggest that "power law" systems explore the available space in such a way as to yield a high diversity for the power law variable. Although the correlation between power law and diversity is purely indicative, it is highly resonant with the main proposition of my work. The topic of diversity constitutes the fourth and final part of the literature review. Its correlation with power law will be discussed in the chapter on theory.

2.5 DIVERSITY

2.5.1 STATING THE PROBLEM

Diversity is a fundamental property of many systems and one of the main drivers of technological innovation. At the same time the topic of diversity is one of the most neglected topics in economics, management and business studies. Although Marshall wrote that *'the tendency to variation is the chief cause of progress'*[262], the rational approach underlying neoclassical economics tends to overlook the often-random contribution offered by diversity. Less orthodox economic schools, as for instance the Austrian school[263] that coalesced around Shumpeter and Hayek, have always valued diversity as a fundamental ingredient in economic systems. More recently, evolutionary economics, and those branches of economics that draw inspiration from evolutionary theory and biological models[264], have stressed similar concepts. For instance Saviotti

---

[262] (Marshall 1916) p.335

[263] (Peters 1999)

[264] (Dosi 1988; Dosi and [et al.] 1988), (Gibbons and Metcalfe 1988)

writes that: "a *trend towards growing variety... is one of the fundamental trends in economic development*"[265]

2.5.2 EVIDENCE FROM HISTORY AND ARCHAEOLOGY

Diversity is inextricably related to innovation and connectivity. To demonstrate my point I will use a few examples from history and archaeology.

The transition from the Palaeolithic to the Neolithic (from tribal hunter-gatherers society to agriculture-based society) was the outcome of two crucial discoveries: the discovery of agriculture and the domestication of animals[266]. In both cases the transition is documented to have taken place in the Middle East (the so-called Fertile Half Moon area - FHM[267]), and independently in China, New Guinea and America[268], between twelve and ten thousand years ago. However, among these areas the development of agriculture and domestication of animals were particularly fast and successful in the FHM area. The evidence provided by (among others) archaeology, ethno-biology, ethno-meteorology, geography and paleo-genetics point out that in the FHM the Palaeolithic-Neolithic transition was facilitated by the higher degree of diversity, in terms of availability of diverse natural resources, with which the hunter-gatherer societies could experiment[269]. The diversity acted thus as a form of absolute advantage. The aspects of diversity were:

1. The FHM constituted the largest extension of contiguous land with a temperate Mediterranean climate.
2. The large availability of land and its internal variety (mountains, costal regions, plateaus, semidesertic regions, etc., but without insurmountable major obstacles for

---

[265] (Saviotti 1996) p. 93

[266] (Diamond 1997)

[267] an area including modern Israel, Lebanon, part of Syria, Jordan, North Iraq, West Iran and Central Turkey

[268] see (Diamond 1997) (Cavalli-Sforza 1995)

[269] The FHM did not have large rivers, forests and extensive access to seas, thereby limiting the stock of resources upon which hunters-gatherers societies survive. Moreover the stock of large mammals had been depleted by centuries of hunting. In the fertile half-moon area the gazelle disappeared at the time of the introduction of agriculture. Therefore hunters-gatherers societies found more and more difficult to support their traditional lifestyle and were in a way forced to invent their way around the depletion of scarce resources.

internal communication) propitiated the emergence of the largest variety of annual species of plants as a response of strong internal climatic variations. The importance of this vegetable abundance cannot be over-estimated. Among the estimated world total of 56 species of herbaceous species with characteristics amenable to agricultural exploitation (that is large seeds and domesticable properties), the HFM's palaeo-farmer could experiment with 32 of them[270]. For contrast, Australia had only two.

3. The FHM showed the highest orographic diversity, from the lowest depression on Earth (the Dead Sea) to high mountains. This made possible to diversify cultivation and harvesting periods and transfer low land species to mountainous areas and vice versa.

4. The FHM also had the highest variety of domesticable large mammals[271]. The candidate progenitor species were diffused unevenly on the globe, with the highest concentration in Eurasia. In this macro-region and in particular in the FHM out of 72 candidates, 13 were domesticated (the total stands at 14) with a percentage of success of 18%. In contrast only one species was domesticated in America, the lama, out of 24 candidates. No species were domesticated in Sub-Saharian Africa (out of 51 candidates).

All these factors point to the fact that certain types of diversity (biological, meteorological, geographic) were instrumental for the onset of the Neolithic age.

But there is something more. The areas of origin of the 4 most important domestic animals (goat, sheep, oxen and pig) and of the 8 progenitors of modern agriculture (spelt, einkorn, barley, lentils, peas, chickpeas, cickling and flax) were scattered all over the FHM and their diffusion areas did not overlap. For instance, the proto-sheep came from the central part of FHM, the goat from the Zagros Mountains (now part of Iraq), the pig from the North, and the oxen from Turkey[272]. Nevertheless the innovations quickly diffused to all FHM. This was because no major geographical barrier prevented communication, such as the Sahara desert, the Alps or the Amazon forest, and in fact the

---

[270] (Diamond 1997) p. 140

[271] (Diamond 1997) p. 162

[272] (Diamond 1997) p. 141

presence of trade is well documented well before the Neolithic. By contrast, the Inca empire managed to domesticate the lama but never invented the wheel, a complementary technology, which was instead available two thousand kilometres north (a comparable distance with the FHM) in the Aztec empire. The lama and the wheel never met because the Amazon forest and the Andes prevented diffusion of innovations.

This point will be central in this work. Diversity (biological, social, economic and technological) is a necessary condition for the innovation process to happen, but rarely sufficient. Without the linkages provided by geography, trades and communication technologies, diversity would result in isolated islands, fragments, maybe endowed with isolated evolutionary pearls, but unable to cross-fertilise and be cross-fertilised. Connectivity and diversity are inextricably linked in a dynamic dance, in which one feeds the other. Linking islands of diversity (that could be described as pools of under-utilised resources) suddenly increases the availability of resources and the potential for discovery of synergistic or symbiotic behaviours, some of which may result in the establishment of new channels of communication, thereby feeding back on connectivity[273].

The non-linear relationship between connectivity and diversity works in both directions towards progress and regress. 15th century imperial China is a striking example: for approximately two thousand years China was at the forefront of technological innovation. The list of innovations that follows is long and not exhaustive: porcelain, trans-oceanic navigation, paper, gunpowder, mobile character for printing, writing, canal locks, magnetic compass, rudder, cast iron, drillings, wheelbarrow, advanced astronomy, sophisticated tools to measure time, etc. In the same century in which Columbus discovered America with three tiny and relatively primitive caravels, China's traded with East Africa by means of ship convoys of hundred of ships that could carry up to twenty eight thousand people[274]. Suddenly this technological blooming

---

[273] The coral reef ecology in the Pacific Ocean is probably the most diverse place on earth because over the past geological eras the changing of the ocean level isolated internal micro-environments where mutation, selection and genetic drift operated on local sub-populations to differentiate via speciation the ecology and zoology of the marine animals and plants. When the water level was raised again, species could migrate away and interact with a newly diversified environment, promoting quick adaptation and the complexification of the ecosphere. We see in the case of Asian coral reef the demonstration that the effect of diversity and connectivity is to build further diversity and more niche complementarity.

[274] (Diamond 1997) p. 412

stopped. Several hypotheses have been advanced (see Mokyr[275] for an exhaustive discussion), spanning from the conservative character of Confucianism to the lack of protein of a rice-based diet. None of them explain China's sudden technological decadence. The most persuasive one has to do with the rigid climate of conservatism that started to dominate the Chinese ruling classes from the 15th century and the strong tradition of political and administrative centralism dominant in such a vast empire. This actively discouraged changes in social, economic and cultural matters. Trans-oceanic navigation was banned, trades with 'barbarians' (non Chinese) discouraged. Although China thrived on internal differences, the political influence of the Court was pervasive. Innovation in China was condemned by the uniformity imposed by a powerful political elite to the point that technological awakening had to wait until the empire collapsed and a more technologically tolerant system took control of China in the 20th century. Although Europe was not exempt from intolerance and aversion to innovation (for instance the condemnation of Galileo by the Catholic church at the beginning of the 17th century isolated Italy from scientific and technological progress for centuries), Europe's internal political and social diversity made possible for prosecuted scientists to find new heavens of tolerance and support by monarchs or powerful elites eager to exploit the advantages of technological innovations. Instead the lack of political diversity exposed China's science and technology to the control of a tiny but powerful political elite. It was (among other factors) the lack of political diversity that condemned China to decadence.

Japan followed a similar destiny. In 1453 two Portuguese adventurers stunned the Japanese by showing the harquebus[276]. A century later, Japanese guns were among the most sophisticated in the world. But, when (following a political decision) Japan became effectively sealed to any external trade and influence with the external world, the use of guns was first socially restricted and then turned into ceremonial objects. Finally they completely disappeared. This is a striking case of technological regress due to restriction in connectivity with the external world (which isolated Japan from the need to compete with the non-Japanese world), and to a rigid climate of social conservatism (as in China), which was favoured by political homogeneity within Japan.

---

[275] (Mokyr 1990)

[276] (Diamond 1997) p. 257

On a smaller scale, the first Tasmanian colonists (who represented the last wave of a migration movement that spread Homo sapiens over all the Indian Ocean) enjoyed a technological level similar to that of the Australian aborigines of the Palaeolithic age[277]. The increase in sea level that followed the end of the last glaciation cut off Tasmania from Australia. The resulting isolation was broken only when the first Europeans arrived on Tasmania in the 17th century. They witnessed the most primitive culture on earth, far below the Palaeolithic level of their ancestors who had colonised Tasmania. The case of Tasmania isolation again shows the consequences of connectivity loss. The geographic and cultural isolation from the rest of the Palaeolithic culture determined a dramatic reduction in the Tasmanian environmental variety. This had a knock on effect on the level of technological complexity that the population already mastered[278]. A more homogeneous environment is not only less innovative than a more diversified one but it is also more at risks to lose its current technological level.

To summarise, the previous examples shows a number of points:

- Diversity is a fundamental dynamic property of social and economic systems.
- A dynamic analysis of diversity can not be done without a parallel analysis of connectivity.
- Loss of connectivity takes the system closer to equilibrium with a dramatic loss of internal diversity, thereby affecting the capability to innovate and threatening the technological status quo.

These points raise a number of issues:

1. Why has the topic of diversity been relatively overlooked in social sciences?
2. What exactly is the relationship between diversity and innovation?
3. If diversity and connectivity are related, what type of relationship links them?
4. Is there a natural trend toward an increase in diversity?

---

277 the capability of manufacturing objects from bones (for fishing), manufacturing nets, boomerang, etc.

278 It seems to be a masterly demonstration of Ashby principle of requisite variety – see section 2.3.1 and a perfect illustration of Maturana and Varela's idea that the environment is a relational property of the entity under consideration, linked with it through the entity's external structure – see section 2.3.4

5. Is there a minimum threshold of diversity and connectivity to ensure an innovative and adaptive system? Are diversity and connectivity subject to critical mass dynamics?
6. Can we measure diversity and connectivity? And if yes, how?
7. And finally, regarding the object of this thesis, can an analysis based on diversity and connectivity help explain the behaviour of geographic industrial clusters?

I will in the following tackle some of the above questions. In this chapter I'll select the relevant sources that will help me address the above questions in the general context of the econosphere.

The first part of this section will describe the relevant literature on diversity, drawing from ecology, economics, innovation theory and focus on the critical relationship between a diverse environment and innovation. In the second part I will examine an important advantage conferred by diversity, that is resilience against the risk of technological lock-in. The third section will provide a context for the previous two aspects introducing the *requisite variety* argument put forward by Ashby. The fourth section is probably the most important: in this section I will discuss the dynamic property of diversity and show that diversity is a natural feature of self-organising systems. I will make large use of complexity theory and in particular of the theory put forward by Stuart Kauffman on autocatalytic sets and by Per Bak on self-organised criticality.

2.5.3 DIVERSITY, INNOVATION AND ADAPTABILITY

The diversity of markets, goods, production techniques and consumer demand has been a important underlying assumption in modern economics. The theories and models of division of labour (Adam Smith, Ricardo's comparative advantage, Marshall's analysis of geographic industrial agglomerations[279] and Schumpeter's *gales of creative destruction*[280] all depend in a way or in another on diversity in markets, goods, consumers and techniques of production. Yet, although central, the analysis of diversity has always remained a rather implicit assumption. A similar fate has the topic of diversity suffered in management and business studies, often mentioned, seldom developed.

---

[279] (Marshall 1919)

It is not in the ambition of this work to discuss the reasons of this benign neglect, but I will nevertheless suggest few reasons (which will be useful in a later part of this work), based on the role of technology and on the Newtonian assumptions of neoclassical economics and management studies.

2.5.3.1 Economics and diversity

Economics is centrally concerned with the process of growth. Mokyr[281] distinguishes between four mechanisms of economic growth: Smithian, Ricardian, Solowian, and Schumpeterian: the first is a function of increase in capital stock (increase in investment), the second is tied to commercial expansion resulting from comparative advantages of trade, the third depends on (roughly speaking) economies of scale and scope and, finally, Schumpeterian growth depends upon the increase in the stock of human knowledge, whose tangible result consists in new technologies and consequent productivity growth. Technological transformation, the engine of Schumpeterian growth, has traditionally been treated as a '*black box*'[282] by mainstream economics theory. The *black box* included anything whose effect was either to modify extant production techniques or to generate new markets (via technological innovation). In either case, the introduction of new technologies was assumed to take place gradually without significantly perturbing the general conditions of economic equilibrium. The worst-case scenario, radical innovation, could be dismissed thanks to three assumptions. First, perturbations inevitably led to a new equilibrium level and constituted an interval between otherwise reigning equilibrium situations; second, the time span of non-equilibrium were short compared to equilibrium time; third, radical innovations were described as exceedingly rare or non-existing at all[283]. The *black box* approach could reasonably neglect major technological transformations.

---

[280] (Schumpeter 1939)

[281] (Mokyr 1990)

[282] (Rosenberg 1982; Rosenberg 1994)

[283] In the most extreme version, radical innovations were simply broken down to a cumulative sum of incremental changes, thereby denying that technological change could upset equilibrium conditions (for a discussion see Mokyr discussion of macro and micro-inventions (Mokyr 1990) and (Rosenberg 1982))

Three great systems of thinking contributed to downplay the importance of technological change. The Darwinian paradigm of gradualism[284] assumed as the driver of evolution and transformation (tradition fully accepted in economic theory) long series of accumulating microscopic changes (therefore not able to perturb equilibrium). The Newtonian method of scientific inquiry concentrated on finding laws explaining the behaviour of objects and was stunningly successful especially in explaining motions under conditions of continuity and equilibrium. Finally, Calculus provided the essential mathematical tools to natural sciences (although not to Darwinism) but with a limitation: the condition of continuity, which underpins calculus, made the use of infinitesimal analysis powerful when the problem could be linearised, usually possible under conditions of equilibrium[285].

Although Newtonian and Darwinian systems of thinking are largely incompatible[286], they both shared the feature of continuity/gradualism. This fundamental assumption shaped the formation of the paradigms under which research in physics and biology was carried out, either at the level of tools/methodologies, or the level of the definition of what constituted an acceptable area of inquiry[287]. By consequence in all the disciplines where the influence of physics and biology was strong – and among others these included economics, sociology, and the more recent management studies –gradualism and equilibrium were *de facto* accepted. Why is this relevant? As Schumpeterian growth (due to radical new technologies) is strongly a non-linear[288] phenomenon, the assumptions of gradualism and continuity do not hold. Shumpeterian growth is a type of swift change, fraught with serendipity, autocatalytic aspects and unintended consequences. These features make the problem of technological change intractable with the tools of the

---

[284] (Gould 1977; Dawkins 1989; Dawkins 1996; Gould 2000; Mayr 2001), see also the very interesting discussion on the origin of *uniformitarianism* in (Raup 1999)

[285] (Prigogine and Stengers 1997)

[286] see 2.5.3.2

[287] see Kuhn for a discussion of this point (Kuhn 1996)

[288] Non-linearity refers to behaviours that cannot be described by linear systems of equations and where the principle of super-position is not valid. Super-position tells us that under certain conditions the results of the actions of different forces can be summed up to describe the overall reaction to a system (Nicolis and Prigogine 1989) pp.58-9. For instance in a crowded room we may still be able to follow different conversations taking place at the same time. In a complex system super-position does not apply because emergent properties can not be reduced to the actions of forces working in isolation from each other.

above-mentioned systems of thought. No surprise that the *black box* escape was very useful.

The second reason of the benign neglect of diversity concerns a methodological and philosophical issue. I will illustrate the point using the General Competitive Equilibrium Theorem proposed by Arrow and Debreu[289], for which they were awarded the Nobel Prize for economics. The General Theorem is one of the most advanced and complicated general models about a modern economy and constitutes the culmination of centuries of research in economics. Arrow and Debreu's theorem basically states that given a set of goods, with all their relationships of substitutability and complementarity and a set of circumstances (initial conditions, to use the language of physics), some of which may be fortuitous, exists a market price for each good, such that the market clears. The theorem assumes complete rationality of the decision-makers and pre-states all the possible set of prices, relationships and circumstances of the economic system. The diversity of the economy enters as a datum, but, because a) the variety of outcomes is pre-stated, b) the agents are perfectly rational and c) they base their decisions on complete information, nowhere in the theory the effects of diversity (that is innovation, evolution and surprise) play any role. Diversity is an input, but its effects are ruled out as outputs. The general competitive equilibrium theory is only an example, but indicates perfectly well that the economists' analytical approach requires, in order to work with the instruments of calculus, the perfect description of outcomes, agents, relationships and type of risks. But reality is not pre-stateble, rationality is bounded and any description is by necessity sub-optimal. The number of possible outcomes is often astronomical and therefore non countable, the possibility of associating probability distributions to outcomes often impossible, thereby denying rational decision-making. These very common conditions known as Knightian uncertainty[290] are the indicators of diversity in action, except when dealing with mature industry during periods of stability!

In short, the focus on gradualism, analytical description and rationality made the topic of diversity not central in neoclassical economics

---

[289] See discussion of GCET in (Kauffman 2000) pp.214-5

[290] (Stirling 1998)

2.5.3.2 Economics of innovation

"...The car comes in and drives the horse out. When the horse goes, so does the smithy, the saddlery, the stable, the harness shop, buggies, and in your West, out goes the Pony Express. But once cars are around, it makes sense to expand the oil industry, build gas stations dotted over the countryside and pave the roads. Once the roads are paved, people start driving all over creation, so motels make sense. What with the speed, traffic lights, traffic cops, traffic courts, and the quiet bribe to get off your parking ticket make their way into the economy and our behaviour pattern"[291].

This quote captures the meaning of the Schumpeterian *gales of creative destruction.* Large waves of technological, economical and social changes are unleashed when a macro-invention[292] triggers a series of cascading effects, which redefines entire industries and the life style of a generation. When this happens, changes are relatively rapid, deep and (at the beginning of the cycle) not cumulative[293]. The picture that emerges from the history of technology is one of turbulence with rapid period of upheaval followed by longer periods of incrementalism. The stream of novelties, coming with the irregular pace of macro and micro innovations, feeds a flow of changes that keeps the economic system out of equilibrium. This flow of changes is of crucial importance. In fact by keeping the economic system under a permanent state of flux, it makes the econosphere similar to the biosphere. Here the appropriate frame of description is antithetical to the Newtonian. The states of the system are not definable a priori, general predictive laws are absent and the flavour of the scientific endeavour takes a narrative character[294]. The nature of evolutionary sciences[295], i.e. any science that accepts the driving role of mutation, selection, reproduction and symbiosis, is to understand in an *ex-post* fashion the dynamic of an idiosyncratic series of events, that is, to tell a story, largely based on anecdotal evidence ("*the survival of the fittest*" is in most cases a tautological[296] explanation). The room for predictability is very limited. Jay Gould[297] claims that if the film of evolution on

---

[291] Brian Arthur cited in (Kauffman 1995) p.279

[292] (Mokyr 1990)

[293] in the socalled fluid phase of the Abernathy-Utterback lifecycle model, alternative fundamental designs are pursued. As most of the designs are imcompatible with one another, due to fundamental differences in the basic techniological solutions, learning across different designs is very limited. The emergence of a dominant design changes the situation.

[294] See for the relationship between pre-statement of possibility space and narrative character of science that ensues (Kauffman 2000) pp134-5

[295] (Mokyr 1991)

[296] (Cavalli-Sforza 1995)

[297] (Gould 1990)

earth was rewound, the result would be dramatically different and we would not be here to tell this story. This quick foray into the differences between evolutionary and Newtonian based epistemologies is relevant to illustrate the different starting points of neoclassical on one hand and evolutionary economics and economics of technology on the other hand.

The concept of equilibrium, central in orthodox economics, comes out upside-down. Its validity is restricted to those relatively rare historical periods, in which stable patterns of economic and social dynamic are prevalent. In other words equilibrium becomes the exception, not the rule

Market-pull and diversity

A short premise on some feature of technological innovation is necessary to introduce the relevance of diversity in innovation. The rational models of innovation emphasise a direct linear relationship between emergence of needs and satisfaction of those by means of the triad invention-innovation-diffusion. The process of change is driven by decisions that originate from market needs and terminates with the introduction of new technologies. This mechanism, known as market-pull model of innovation, received strong backing from several empirical studies[298] and was widely believed to be the dominant framework to explain the innovation process. The introduction of new technologies (processes and products) responds to a rational chain of environmental scanning and decision-making which start from the identification of market needs, continues with a decision whether to invest resources in the potential market, prosecute with project-managing the new products development process, and ends with success or failure in the market place. Innovations, piloted by central agencies, military organisations and other institutions usually follow a similar approach, with strategic missions replacing market needs[299]. What is relevant for our discussion is that this approach acts as a powerful filter, selecting among the multiple environmental signals those that resonate with the mental frameworks and paradigms of the people in charge of the scanning and decision (usually marketers and technocrats). Unclear, confused and

---

[298] see in (Rosenberg 1982) the essay on: "*the influence of market demand upon innovation: a critical review of some recent empirical studies*"

[299] (Martin 1994)

fuzzy signals are rejected either because they are not perceived (here the lesson of Kuhnian storicism[300], Weick's[301] sense-making and Cohen and Levinthal's absorptive knowledge[302] are useful), look too risky, or don't seem to have enough potential. The innovations that pass through the filter are usually low risk, incremental, and relatively homogenous with the technologies and products from which they derive. For the same reasons highly radical and disruptive innovation, the ones that open new technological trajectories and new markets, are discarded. What are the consequences of this dynamics with respect to diversity in the econosphere? The continuity that results from the risk-averse approach favours the research of incremental variations within the envelope of the established technological and market trajectories, as these are the only ones that allow probabilistic risk assessment and technology forecasting (typical management tools to assess and control risk). In the language that will be introduced in section 3.1 of the methodology chapter this implies an increase in variety, that is an increase in the number of related species or artefacts, based on the same general architecture[303] or dominant design[304], but not an increase in the number of technological trajectories or market sectors. Market-pull innovation then contributes to only one of the dimensions of diversity.

Technology-push and diversity

An alternative model of innovation is the technology-push model. This model predicates that the origin of new technology does not occur as a response to market stimuli but is more akin to Darwinian mutations taking place in an environment fraught

---

[300] *The history of astronomy provides many other examples of paradigm-induced changes in scientific perceptions. Can it conceivably be an accident, for example, that Western astronomers first saw change in the previously immutable heavens during the half century after Copernicus, new paradigm was first proposed? The Chinese, whose cosmological beliefs did not preclude celestial change, had recorded the appearance of many new stars in the heavens at a much earlier date. Also, without the aid of a telescope, the Chinese had systematically recorded the appearance of sunspots centuries before these were seen by Galileo and his contemporaries. ....*

The very ease and rapidity of with which astronomers saw new things when looking at old objects with old instruments may make us wish to say that, after Copernicus, astronomers lived in a different world. In any case, their research responded as though that were the case. Kuhn: "The structure of scientific revolutions", p.117

[301] (Weick 1995)

[302] (Cohen and Levinthal 1990)

[303] (Henderson and Clark 1990) see also (Sanchez and Collins 2001)

[304] (Abernathy and Utterback 1978)

with Knightian uncertainty and ignorance[305]. In Darwinian evolution mutations are random events upon which selectionist pressures act to filter useful mutations from harmful ones. However, the analogy with Darwinist technology-push against the more Lamarckian market-pull[306] is of limited utility if one accepts in full one of the tenets of Neodarwinism: that is, according to the standard model of evolution[307], natural selection is largely deterministic. This is in contrast with the irreducible character of chance in technological evolution, which operates not only at the level of supply (mutations) but also at the level of adoption (the process that should be governed by the analogous of natural selection). Both Lamarckian and NeoDarwinist analysis are ineffectual to explain the data. Technology-push, radical innovations are events that prop up unguided by market forces and often unrelated with market forces. The case of the QWERTY keyboard[308], and of many other technologies that contributed to open new technological trajectories, stand to demonstrate the ineffectiveness of explanations based on determinism. Why is this relevant for our discussion on diversity? At least for two aspects. First, technology-push innovations produce generic solutions that can potentially be applied to a disparate array of fields. Second, the undirected aspect of technology-push innovation results more often than in the market-pull case in radical innovation, thereby opening new markets where there was none. In terms of diversity, technology-push innovation gives a two-fold contribution: first, by diversifying the pool of potentially usable technologies within a market sector and second by enlarging the number of market sectors.

Market-pull innovation responds to planning, forecasting and management techniques. Technology-push does not. An example will help make the point. 3M Post-it innovation was the unintended consequence of the 'wrong' invention, one that yielded an extremely weak and impermanent adhesive instead of the strong one that was sought[309]. The invention was communicated within 3M but remained dormant until a scientist, Art Fry,

---

[305] The distinction with the two has to do with outcomes and probabilities of outcomes. In knightian uncertainty outcomes are predictable but probability patterns are not. In a situation of ignorance neither outcomes nor probability distribution are describable (Stirling 1998)

[306] (Martin 1994)

[307] (Dawkins 1989; Dawkins 1996; Mayr 2001)

[308] (Arthur; David 1985) (Gould 1997)

[309] (Gundling 2000)

realised that it could be useful to mark the pages at his church choir practice. During the product development (that occurred within the protecting environment of the 15% rule[310]), Art's group expanded the application range from bookmark to note paper. In spite of this, the new product was not given the green light for commercialisation. At that point the developing group resorted to guerrilla tactics and distributed the products internally to generate enough demand and behavioural change in order to get the product commercialised. The result was an outstanding success. The short example provides some instructive lessons: first, innovation is serendipitous and no amount of planning can ever imagine to achieve the same results via a strategic rational approach; second, applications are found during the development and commercialisation phase at the interface between the developing team and the final users; third, the history of innovation is largely one of unintended consequences. Most of the market-shaping innovations were developed with a completely different market and range of applications in mind[311]. Successful markets are not 'found' (even allowing for a trial and error approach), as if they were islands to be discovered by an entrepreneurial scientist, but are literally developed at the interface between supply and demand forces[312].

The implications of these aspects are drawn in the following section.

2.5.3.3 Exaptations

Let me start with an example. Reasonable evidence[313] points out that biological development of feathers was not driven by a selectionist pressure to develop flight and therefore escape predators, as a traditional reading of evolutionary theory would claim, but to improve thermal isolation. Feathered flying evolved later in evolutionary history. This evolutionary pattern seems to be common. Even language[314] seems to be the result of evolutionary changes[315], which occurred for completely unrelated reasons to language

---

[310] (Gundling 2000)

[311] (Levinson 1997)

[312] (Lynn, Morone et al. 1996)

[313] See (Mayr 2001) (Gould and Vrba 1982)

[314] (Tattersall 2001)

[315] anatomical changes concerning the distance between larynx and pharynx

itself[316]. Evolution is not only blind, but crucial aspects of it escape the allegedly inescapable logic of natural selection. Orthodox evolutionary theory claims that if a random genetic mutation is favourable to the organism's survival, the mutation is selected and passed on to its offsprings. If unfavourable, the organisms do not survive to transmit the mutation. Many mutations, however, are neither favourable nor unfavourable, they are neutral and don't affect the bearer's survival chance in the environment. Under some circumstances (for instance genetic drift in small populations[317]) the mutation may spread in the population and lay dormant in the genotype, until either environmental changes give an environmental meaning to the dormant 'capability', or an endogenous dynamics of development between the dormant feature and the rest of the phenotype takes place invisibly outside the sphere of control of natural selection. At that point the speed of evolution accelerates and fundamental changes seem to happen swiftly. This type of evolutionary dynamics is known as exaptation[318].

The point I want to make is that the mechanism of exaptation is useful to explain technological evolution. The elements that characterise the Post-it innovation story, that is, serendipity, emergence of new applications during the development phase and unintended consequences, make more sense when interpreted as an exaptation. Technology-push mechanisms provide a source of random mutations under the form of new technologies that are often stored in the organisational memory. In occasion of some sort of environmental change (in a coevolutionary way), the stored technology (or mutation) is rediscovered and applied into a new context and a set of applications developed into a new market niche by users and developers in a sort of conversational feedback loop. I will develop the implication of exaptation in section 3.2.4.

---

[316] The emergence of photosynthetic life seems to obey the same pattern: Lynn Margoulis[316] supports the popular Gaia (Margulis 1998) hypothesis showing that the appearance of conditions (mix of oxygen and methane gases) favourable to photosynthesis-based life was a bioproduct of the activity of cyanobacteria, for which oxygen was poisonous and therefore excreted.

[317] see (Cavalli-Sforza 1995) pp. 97-102

[318] The concept was introduced by Gould and Vrba (Gould and Vrba 1982). See also (Tattersall 2001) (Kauffman 2000), (Mayr 2001) p.207

2.5.3.4 The advantages and cost of diversity

I will in this section review some evidence coming from the literature relating diversity and economics. I will integrate the excellent online paper by Stirling[319] with material coming from other sources.

Although diversity underlies economics (diversity of consumer tastes, production techniques, products and technologies are fundamental for competitive theory) and without which no economic system would exist, much economic discourse makes some simplifying assumptions that reduces that diversity. First, the *homo economicus* (the rational man), which represents the other end of the spectrum from the *homo emoticus* (the emotional man)[320], covers with an assumption of rationality the diversity of agents' behaviours. Second, the *black box*[321] assumption, by eliminating the problem of how technologies emerge, compete and come to be used in production, excludes the main source of diversity increase, i.e. technological change. Third, the tacit assumption of gradualism and linearity in evolutionary change also operates as a simplification, as it tends to favour evolutionary technological evolution (a succession of small modifications that by accumulation transforms technology A in technology B) over more radical ideas concerning the revolutionary aspects of the origin of new technological species (based on bifurcation resulting in an extreme case, in the rapid increase in the number of fundamentally different set of alternative designs).

Evolutionary economics[322] refocuses the attention on the relationship between economic growth and technological change and consequently on the causes of technological change. The empirical observations coming from this field of study make economic diversity dependent on technological innovation and explain the diversity's advantage in terms of resilience against uncertainty. In other words, economic diversity enhances adaptability to turbulent environments by acting as a resource pool from which suitable strategies or new technologies can be extracted to suit a fast changing unpredictable world. This insight is well known in finance management: managing a portfolio of diversified options minimises risk and optimises survival rate. This is true

---

[319] (Stirling 1998)

[320] (Sigmund, Fehr et al. 2002)

[321] (Rosenberg 1982; Rosenberg 1994)

[322] (Nelson and Winter 1982; Dosi and Orsenigo 1985)

also in management of technology: a diversified pool of technologies and projects incubating in R&D increases survival rate by rising the probability of hitting the right technology at each innovation phase. However the very important point to be made concerns the causal link between diversity and innovation (and consequently economic growth). Gibbon and Metcalfe[323] notice that innovation in several areas seem to be correlated to "*the degree of economic variety contained within it*". Saviotti[324] in several papers and books dedicated to the role of diversity in economic growth states that the "*a trend towards growing variety is ... one of the fundamental trends in economic development*"[325]. Saviotti sees technological diversity as playing a fundamental and necessary role in maintaining economic growth.

An interesting observation[326] stresses the still unexplained fact that most of the innovation activities (for instance as measured by patenting) take place in relatively few places around the earth (for example the computer is largely a Californian history[327]). The role of laboratories such as the Bell, Palo Alto Xerox, the IBMs, universities such as MIT, Berkeley, Oxbridge, and location such as Silicon Valley, etc accounts for large part of technological innovation in the world. What characterises this geographic concentration is the spread of activities within them, that is their internal diversity. A parallel observation is made by Scheinkman[328] when he states that "*growth of cities appears most strongly correlated with industrial diversity and not with concentration within single industries*". The evidence above indicates correlation between economic diversity and innovation rather than causality. Contributions which instead frame the relationship between diversity and innovation in terms of causality are listed below:

1. Complexity theory[329] has elaborated a sophisticated framework connecting micro-diversity of agents with growth, albeit not explicitly economic, via the concepts of

---

[323] (Gibbons and Metcalfe 1988)

[324] (Saviotti and Mani 1995; Saviotti 1996)

[325] (Saviotti 1996) p.93

[326] (Storper 1997; Castells 2000)

[327] (Cringely 1996; Castells 2000)

[328] cited in (Kauffman 1995) p.295. A similar point is made in the discussion regarding urbanisation vs. localisation economies; see section 7.2.1

[329] (Holland 1995; Kauffman 1995; Arthur 1996; Kauffman 2000; Allen 2001). The contribution of complexity theory to the dynamics of diversity is introduced in sections 2.3 and 2.5.4.

self-organisation, emergent properties and auto-catalytic cycles. These are discussed in section 2.3.

2. On a resonant line but from a sociological viewpoint, the work of Callon[330] on techno-economic networks stresses the importance of fostering linkages between diverse actors. In Callon's view, innovation is a by-product of the conversations taking place between carriers of different knowledge who have to 'translate' their language into the network context.

3. This type of discussion blends well with the large and rapidly expanding field regarding the network organisation. Networks are seen as a natural form where internally coherent and highly connected local pockets can coexists and interact with other pockets. The weaker links between pockets ensures that pockets maintain their identity and coherence whilst, at the same time, making possible the interaction of diversity. Networks are in the words of Grabher the place where "*the diversity of organisation*" self-organises into "*the organisation of diversity*"[331].

4. Micro-exploration or "*error making*"[332] as Peter Allen points out is the property of non-average agents to carry out parallel exploration of their surroundings. The degree of innovativeness of a system comes to depend on the number of idiosyncratic agents and on the specific trade-off between energy-consuming explorations and energy-producing exploitation of resources.

If diversity can be seen as providing important advantages to aggregated economic actors, however, diversity comes with a cost and as Weizman writes: "*the laws of economics also apply to diversity ...There are no free lunches for diversity*". The costs of diversity (or diversification policy) are essentially lost economies of scale via standardisation (the opposite of diversity) and an increase in transaction costs due to the need of finding and coordinating a host of different and idiosyncratic actors. The contrast between 'centrifugal' forces that increase internal diversity (to maximise innovation rate) and centripetal ones that increase homogeneity (to maximise economies of scale and minimise transaction costs) leads to the research of an optimal trade-off between

---

[330] (Callon 1991)

[331] (Grabher and Stark 1997)

[332] (Allen 2001; McKelvey 2001)

diversity and standardisation[333]. Stirling correctly points out that achieving the trade-off is problematic as diversity appears on both sides of the equation: for instance if economies of scale are promoted by a decrease in diversity, economies of scope[334] are improved by an increase in diversity. In any case there seems to be an overall trend toward shifting the trade-off towards an increase in overall diversity due to a series of factors, the most important of which being the ICT revolution[335]

2.5.3.5 Ashby law and diversity as a response to ignorance

Ashby's law of requisite variety[336] predicates that a resilient system is characterised by the capability to match environmental stimuli (or shocks) with a correspondent type of effective responses. There is therefore a set-to-set type of correspondence by which the system can put in place a specific type of strategy (I use the term strategy in the meaning proposed by Axelrod[337], that is specific and predetermined response) to fight or match an environmental threat or opportunity. Although the original formulation referred to cybernetic, a recent example provided by Kauffman[338] will clarify the concept. Szostack and Ellington demonstrated that, given a library of RNA molecules 100 nucleotides long, the probability of binding a RNA molecule to any small organic molecule is approximately one in hundred million (10^{-8}). This means that a set of RNA molecule (*a universal RNA toolkit*[339]) 100 millions rich is able to stereo specifically attach and potentially neutralise (or catalyse) any small organic molecule (target). It is then not by magic that the approximate length of human immune system is 100 millions. This number represents the requisite variety, by which the immune system is able to devise a strategy to recognise and attack any intruder in the system.

However the requisite variety concept rests on a critical assumption to be operationalised: responding to environmental stimuli requires a specific capability in

---

[333] From a absorptive knowledge standpoint an increase in diversity beyond a certain limit would cause the loss of a common language and grammair (see (Maskell 2001))

[334] (Bellandi 1995)

[335] (Malone and Laubacher 1998) (Castells 2000)

[336] (Ashby 1956). See section 2.3.1

[337] (Axelrod 1984)

[338] (Kauffman 2000) pp.27-28

[339] the expression is by Kauffman

order, first, to recognise a new stimulus as different from what experienced in the past, second, to assess its structure, and third to mobilise (if present) or manufacture (if absent) a response. However, as Kauffman's example shows, the set-to-set correspondence implies that the number of stimuli is not only limited but also computable. No system will ever be able to match a limited but astronomical set of stimuli, at least without sinking into chaos. Here is where we find a problem. To introduce this point I need to spend few words on the concept of risk, incertitude, and ignorance[340].

Let us imagine we are a system planning the set of responses to future stimuli. It could be a nation state forecasting the next terrorist threat in order to plan an adequate countermeasure or a company deciding the allocation of resources to new product development projects on the basis of the hypothesised competitors' strategies and/or new entrants' moves. Our company will use reasonable models based roughly on two variables: knowledge about outcomes and knowledge about likelihoods of the outcomes. The former refers to the probability distribution of a certain outcome and the former to the number of possible distinct outcomes. We run into the following situations:

First, the situation under analysis is amenable to a limited (not astronomical) number of outcomes and these are describable by means of probability distributions. This occurrence is described by the term risk. Second, even though probability distributions can be attached to outcomes, the latter are poorly defined or non computable (astronomical number). Fuzzy logic may help when outcomes are overlapping. The third case occurs when outcomes are definable but probability distributions are not, that is, it is impossible to speculate about which events are more likely to occur. This situation was described by Knight as uncertainty. Finally when neither outcomes nor probability distribution are assessable, we are in the realm of ignorance. We simply *don't know what we don't know*[341].

[340] An excellent introduction is in (Peters 1999) and (Stirling 1998)

[341] Curiously and ironically as (Laing 1970) points out this situation gives rise to expertise:

"*There is something I don't know*
that I am supposed to know.
I don't know what I am supposed to know,
and yet am supposed to know,
and I feel I look stupid

An Ashby strategy is applicable and appropriate only in the first case. It is still valid in the second and third, albeit with different levels of difficulty, but is entirely inadequate with the fourth. What is the point of preparing defences against a totally unknown enemy?

Interestingly technological change seems often to fall into the ignorance camp[342]. Both the anecdotal evidence and the academic analysis suggest that major technological changes (macro-inventions in Mokyr's parlance[343]) unleash a long lasting and pervasive series of *unintended consequences*[344] that create entire new sectors or reshape existing ones. Moreover the value and utility of many innovations go almost invariably beyond the intentions of the inventors (or manufacturers). The examples are uncountable. For instance the whole field of environmental impact of the nuclear[345], the history of the Internet[346] and several other examples show that the patterns of usage of new technologies develop only after the technology becomes available for experimentation. Also, the process of research, development and adaptations of new radical technologies requires the elaboration of a new paradigm that rarely if ever precedes the technology.

---

if I seem both not to know it
and not know what it is I don't know.
Therefore I pretend to know it.
This is nerve-racking
since I don't know what I must pretend to know.
Therefore I pretend to know everything.

I feel you know what I am supposed to know
but you can't tell me what it is
because you don't know that I don't know what it is

You may know what I don't know, but not
that I don't know it,
and I can't tell you. So you will have to tell me everything".

342 The role of uncertainty as a necessary condition for competition has been explored by the Austrian school of economics and in particular by Hayek. Hayek wrote: "*Competition means decentralised planning by many separate persons*";and "*all economic problems are caused by unforeseen change which requires adaptation*" (Quoted in (Peters 1999) p.96 and p.115). The market is seen by Hayek as a structure that results from the intermediation of local spheres of personal uncertainties. Consequently uncertainty comes to have a fondamental role in economics. It is the unavoidable element of uncertainty that, according to Hayek (writing in 1946), would have condemned the socialist experiment to failure.

343 (Mokyr 1990)

344 (Levinson 1997)

345 (Pool 1997)

The profound impact of radical new technologies has been historically framed by Kuhn[347] as paradigm shifting. The long wave of unintended consequences deriving from often seemingly small inventions (many fundamental inventions from the spinning loom[348] to the automobile, from the telephone[349], to the internet[350] went virtually unnoticed) display their effects by transforming existing patterns of usage and by introducing new applications that transform the social habits related to those technologies. The introduction of the Internet and the convergence between Internet and mobile telephony have not only changed the way people communicate from homes and offices but also transformed *tout-court* the concept of communication. The paradigm associated with distance and time has been slashed (distance is no obstacle to digital connectivity); the progressive transformation of the handset into a microcomputer is changing the perception we have of the telephone and blurring the distinction between the computer (traditionally associated with work and number crunching) and telephone (associated with voice and communication). Because communication is an integral aspect of almost all human activities, a fundamental change in communication channels has affected and will affect (to an extent that is impossible to foresee now) all activities that depend on communication. Arpanet, the first prototype of the Internet appeared in 1969 under the auspices of DARPA[351], escaped rapidly from the control of its military sponsors. There are several reasons why. One of them was that the network was used to establish an inter-university network of science-fiction lovers to such an extent that DARPA was forced to separate Arpanet (military) from the Internet (civilian). The fact that the Internet would spur a business revolution (e-business), a radical change in organisation structures (virtual organisations, ecommerce, an alternative to product development, such as the Open Source movement, e-bay, etc.), the *death of distance*[352] and a political erosions of the nation state's control over information was totally outside the scientists' and sponsors' minds.

---

[346] (Cringely 1996)

[347] (Kuhn 1996)

[348] (Mokyr 1990)

[349] (Levinson 1997)

[350] (Gleick 2002)

[351] (Castells 2000) p.45

Ignorance is therefore inextricably linked to any innovative activity[353]. It is not to be intended as negative (or anyway not necessarily) or as a residual (Stirling) of what even a powerful mind can or can not anticipate. Paradoxically, ignorance increases with the amount of knowledge. Like in an expanding sphere, where the surface increases with the volume, the interface between the known and the unknown increases with the accumulation of knowledge. Ignorance is first, a necessary feature of the exploration associated to venturing in unknown waters, and second, of the fact that the structures (social, economic and mental) which allow agents to make sense and to operate '*without trembling hands*'[354] are themselves an effect of the co-evolutionary dance between agents and environment. As such, they can not pre-date the environmental change that is triggered by a technological innovation.

If this is the case, then, Ashby law of requisite variety does not work. The impossibility of measuring the environmental diversity (and therefore of adapting the span of system's response to the variety of environmental stimuli) makes an Ashby strategy ineffectual. Allen[355] has suggested a modification of Ashby law called *law of excessive variety*. Allen states that a system has to strive for an excess variety in order to compensate for (what we can call) the *ontological* ignorance we have of the environment. Diversity becomes then an effective way to counteract the unpredictability of the environment under conditions of ignorance and uncertainty. The question of how to diversify and to what extent remains however open. The power of Ashby theorem resided in the computability of the system-environment dynamics. Once that is lost, because ignorance is an intrinsic property of exploration, we have to resort to strategies that are potentially able to cope with the dimension of ignorance, that is, diversity.

I will briefly mention a further advantage of diversity, which will be explored in more details in the section on dynamic diversity. The real reason why diversity works in a systemic context is not so much the point that a larger response pool raises the probability of having the right strategy in stock. This is true but in presence of ignorance

---

[352] (Cairncross 1998)

[353] As Hayek wrote: "*The solution of the economic problem of society is ... always a voyage of exploration into the unknown*" quoted in (Peters 1999) p.59

[354] (Kauffman 2000) p.160

[355] (Allen 2001)

it constitutes a weak defence. The reason lies in the fact that systems harbouring high internal diversity are better able to show emerging properties and self-organisation, that is, to build a new internal architecture around the new environmental constraints. This is often achieved by reconfiguring the patterns of linkages among modules, attaining in this way a new functional structure.

2.5.3.6 Diversity and the risk of "locking-in"

In this section I will discuss a further advantage of diversity, which has to do with avoiding lock-in. In order to do so, I will first briefly discuss the concepts of increasing returns, autocatalytic dynamics and lock-in. I will then pass to illustrate why diversity is useful to prevent or mitigate the effects of lock-in.

Although the concept of increasing returns (or positive feedback) is not new in economics, its importance has only recently been recognised[356], mainly through the works of Arthur[357] and, in parallel, by a stream of economists, among whom the economic geographers have played a particularly important role. They have been concerned with the problem of the patterns of localisation of economic activities and the stunning asymmetry in its concentration. Fujita, Venables and Krugman[358] exemplify the point:

" ... *the basic problem with doing theory in economic geography has always been the observation that any sensible story about regional and urban development must hinge crucially on the role of increasing returns, ..., the spectacular concentration of particular industries in Silicon Valleys and Hollywoods is surely the result not of inherent differences between locations but of some set of cumulative processes, necessarily involving some form of increasing returns, whereby geographic concentration can be self-reinforcing*".

This position runs counter to orthodoxy in economics, where equilibrium is the condition of reference. Centripetal forces, that takes the economic system (or part of it) away from equilibrium (and that are produced for instance by increasing returns mechanisms such as economies of scale) are treated as fluctuations, or in other words as

---

[356] Sraffa, a Keynesian economist, wrote a paper to demonstrate that increasing returns, if accepted, would have destroyed classical economics (Sraffa 1960)

[357] (Arthur 1996) (Arthur, Durlauf et al. 1997)

[358] (Fujita, Venables et al. 1999) p.2

temporary phenomena, dampened when centrifugal forces (i.e. diminishing returns) take the lead.

Increasing returns have received particular attention since the network economy[359] has turned to the limelight. A highly interconnected and diversified economy is more subject to network externalities[360], whereby a form of self-reinforcing dynamics between economic complementarities causes an increase in the consumption of economic option B as a result of increased consumption of its complementary option A. The classical example is VCR and videotape[361]: as Betamax and VHS were incompatible designs, an early lead in VHS diffusion would encourage consumers and retailers to invest in the VHS camp, creating further demand for VHS's VCRs, in this way stimulating content producers to produce in VHS format. Retailers, content producers and consumers investments would feed back on VHS dominance, creating a self-sustaining dynamics, whereby the attractiveness of a product depends on the number of consumers that choose that product/complementary artefacts and the number of consumers making the fatal choice depends on the attractiveness of the product. In simple terms then increasing returns stand for a situation where the more an option is selected, the more attractive it becomes.

Why are we concerned with increasing returns? Because, by taking the system out of equilibrium, increasing returns create the conditions for specific sectors, defined around a set of technologies and related products, to generate via a path-dependent mechanism a technological trajectory. As there are several ways of taking a system out of equilibrium, the trajectory taken is only one of the possible. Betamax could have won or, if the game had had a slower evolution, we would probably have witnessed a VCR industry with two or more players, grouped around different standards with different geographic dominance. This is for instance the situation of mobile telephony where GSM dominates around 65% of the world market with different standards (and players) operating in America and Japan. Seemingly minor events can play a fundamental role when the

---

[359] (Evans and Wurster 1997; Malone and Laubacher 1998; Evans and Wurster 1999; Castells 2000)

[360] (Teece 1987; Cusumano, Mylonadis et al.)

[361] (Cusumano, Mylonadis et al. 1997)

system is close to a bifurcation point[362], from which several trajectories may start.
The dominance of the internal combustion engine over the steam engine in the automotive sector or of the Stephenson locomotive over the competitors[363] seem to have been caused by contingent events. These facts cast serious doubts about the claimed capability of markets to select the optimal technology given a set of available technologies and market demands[364]. Once a technological trajectory has been started, and the corresponding paradigm developed, it is very difficult or virtually impossible to jump back onto the other branches of the bifurcation. The momentum associated with the positive feedback of increasing returns and even more strongly with its autocatalytic dynamics, and the cognitive aspects of the new paradigm influencing the sense making activity (and therefore the decision making), generate a *lock-in* around the winner technology. The case of the QWERTY keyboard[365] is exemplary. The QWERTY design was a contingent event that was historically frozen by the positive feedback dynamics between diffusion and attractiveness of a standard[366].

The lock-in phenomenon occurs in two stages: first, the action of positive feedback loops determines the emergence of a *winner takes all* situation, and second, the high cost of switching a critical mass of users toward an alternative design congeals the locked in situation.

Lock-ins are not without advantages. Standardisation is one. The emergence of a dominant design[367] after an evolutionary struggle between alternative designs carries the promise of rapid improvement in the selected design due to the massive commitment of resources around it, with the consequent standardisation effects. At the same time, as the lesson of the QWERTY teaches, there are risks. Once universally adopted, a dominant

---

[362] Nobel Prize Ilya Prigogine highlights how even in very simple physical systems (Bénard cells) random events play a crucial role in pushing the system toward one of the branches of the bifurcation (Nicolis and Prigogine 1989)

[363] (Pain 2000)

[364] Natural selection is also claimed to be effective in selecting the best fitness improving mutations (selection of the fittest)

[365] (David 1985; Gould 1997)

[366] It is paradoxical that we still use a design which was expressively designed to slow down typist, thereby representing the worst design for the keyboard industry, after the specific context for which the QWERTY was developed ceased to be important.

[367] see the Abernathy-Utterback model of technological change (Abernathy and Utterback 1978)

design becomes sticky. Dominant design may become associated with group-think and mono-cultures and in some cases with monopolistic positions (the example of Microsoft's virtual monopoly in PC operating systems stands above all others) that distort markets and stifle innovation.

Diversity acts as an antidote to prevent the ossification of the industrial model and of its technological base around a single dominant design[368]. A diverse environment offers more resistance to lock-in because the agents' micro-diversity is more likely to show a distribution of behavioural reactions to the forces causing the lock-in, thereby making more difficult for the lock-in dynamic to take over the whole system. Internal diversity, as I showed in the imaginary history of Meadow and I will expand in the next chapter, preserves the capability of the system to explore its environment, not only by incremental innovation, which takes place within the envelope of the dominant design, but also via radical and architectural innovation, which instead challenge the dominant design by opening new trajectories of development. The architectural element is especially important. In fact, high internal diversity and connectivity breed modularity and the capability for reconfiguration of weakly linked modules. At the same time internal diversity is also likely to result in the mutual dependence between several positive feedback loops. Because a diversified environment is highly internally structured, lock-in dynamics are likely to a) produce dominant design dominance at a niche level more than at a system level and b) generate more mutualism between positive feedback loops.

2.5.4 DYNAMIC DIVERSITY

I have in the previous section shown that diversity (however we may choose to define it)[369] is an essential component of socio-economic systems, is strictly correlated with innovation and the maintenance of vibrant societies, constitutes a useful antidote against lock-in, and bears a cost. I have also shown that diversity is a useful parameter in the discussion of the "*law of requisite variety*". There is however a further aspect of diversity that deserves attention, which concerns the essential role that diversity plays in the making of many systems.

---

[368] a similar point is made in the introduction (section 1.2) when the consequences of the (imaginary) discovery of casein are examined

I will start with an example taken from Kauffman[370]. Kauffman takes the move from the famous Hoyle and Wickramasinghe's[371] observation that the spontaneous generation of life on the earth is so improbable to be in effect practically impossible. Their reasoning is mainly combinatorial. The space of possibilities for instance of combining simple chemical elements into complex biological molecules is so astronomically large and the percentage of useful combinations for life to arise so minuscule, that in the four billion years (age of the earth) available for experimentation nature could explore only a tiny fraction of the possibility space, too small to generate any credible scenario for the start of life. Kauffman's approach is different: if we imagine an initial *primordial soup* of chemical elements, such as that found in inert planetary bodies, the distribution of chemical species therein contained must have allowed (albeit at an exceedingly slow rate) the synthesis of new more complex chemical compounds. It is also highly probable that a percentage of the chemical elements have some catalysing properties, that is, they accelerate reactions. Because the increase in chemical components is rapidly offset by the increase in potential channels of reaction among the elements, given proper environmental conditions and time, the *soup* would develop a broader variety of components, so that the probability that some of the catalysed reactions could achieve closure and generate a giant autocatalytic loop would approach certainty. In the discussion regarding the properties of autocatalytic loop in section 2.3.3, I show that autocatalytic loops behave as seeds of organisations by transforming aggregates into coherent wholes. The emergence of life, or more generally, the emergence of self-organising systems, comes then to depend on the generative property of diversity to promote further diversity with the consequent higher probability that some of that diversity will self-organise into large hypercycle and give rise to systems[372]. The paradox highlighted by Hoyle and Wickramasinghe disappears when the development of order is considered from the point of view of *becoming* more than *being*.

---

369 In section 4.3.1 I will discuss the problem of defining diversity

370 (Kauffman 2000)

371 Cited in kauffman pp.45-6. This is also known as the Boing 747 paradox. The chance of forming a reproductive enzyme are equal to "*the chances that a tornado sweeping through a junkyard might assemble a Boing 747 from the material therein*".

372 In practice the emergence of life is not only much more probable in the scale of available time than the Hoyle and Wickramasinghe's hypothesis would led us to believe, but also more probable than NeoDarwinists would admit.

Kauffman makes a second important point. Above a certain threshold the increase in diversity becomes explosive (as in a runaway reaction). Below the same threshold instead the increase in diversity is slow but gradual. However in biological terms, a runaway reaction implies the annihilation of any organism[373]. In fact, a chemical substance that entered the membrane of a cell would very likely, if successful in generating new chemical substances, disrupt the organisation of the cell by poisoning it, and probably trigger a chain of related extinctions. Homeostasis is instead the rule. How do organisms defend themselves against a runaway increase in diversity? Kauffman suggests they do so by creating boundaries, that is, by surrounding a set of elements and their dynamic reactions with a membrane, thereby filtering the new elements which could give rise to the runaway dynamic. The final result is that cells evolves to a specific threshold, situated just below the threshold between sub-critical and supracritical regimes. The double pressure, between increase in diversity (when in condition of subcriticality) and reduction in diversity (when supracritical) due to negative selection, situates cells and organism on the boundary between subcriticality and supracriticality. This dynamic situation between the order of subcriticality and the chaos of supracriticality suggests an *edge of chaos* frontier, whereby a system can exist in a chaotic, statically ordered and mixed mode. When the system errs into the supracritical region, the extinction mechanism tends to bring it back toward the edge of chaos. Once there the system will oscillate around the *edge of chaos* border by the actions of extinction and autocatalytic reactions. It is highly likely that the mechanisms for reducing excess diversity will work according to a power law distribution. At the same time, Kauffman claims, the biosphere is clearly supracritical[374]. A rather similar situation is found within the econosphere[375], where the relative stability of organisations is matched by an increasing variety and diversity of organisations and players. This point is rather ambiguous. It isn't clear what mechanisms maintain the basic unit at a slightly subcritical

---

[373] Cancer is a form of runaway reaction

[374] The generation of the approximately 7-40 million biological species (the number is highly uncertain, however 7 millions is a lower boundary estimate – see (Raup 1999), pp.47-8, (Mayr 2001) pp.161-3) is witness to the supracriticality of the biosphere.

[375] Some scholars ((Gould 1997)) warns about the extension of *biological* ideas to different fields (for instance econosphere of history of technologies). Others (including (Mokyr 1998), and (Kauffman 1995) do not see a contradiction in it.

level, an assembly of basic units (a group of cells for instance) at the boundary (*edge of chaos*), but the whole biosphere at a supracritical level.

A better approach is taken in Kauffman's 2000 book: Investigations[376]. Kauffman gives an unambiguous definition of the most fundamental unit of analysis of a complex system, the agent. An agent is a unit based on an autocatalytic cycle and able to carry out some thermodynamic work. This is important because, as introduced in section 2.3.5 (on morphogenetic properties of constraints), work requires the creation of constraints in order to spot a gradient of energy (or in other words a disequilibrium situation) and to exploit it. The two properties of an autonomous agent (catalytic closure that provides the dynamic principle of organisation around a set of processes and capacity to carry out work that implies the pre-constructions of tools - which act as enabling constraints - to intervene in the world) ensure first, that the agent is stable (due to the properties of amplification and selection of auto-catalytic loops), and second, that it can actively find sources of energy to survive. At this point we have all the basic blocks to introduce Kauffman's four laws of complex systems[377]:

1. *" Communities of autonomous agents will evolve to the dynamic "edge of chaos" within and between members of the community, thereby simultaneously achieving an optimal coarse graining of each agent's world that maximises the capacity of the each agent to discriminate and act without trembling hands"*

2. *A coassembling community of agents, on a short timescale with respect to coevolution, will assemble to a self-organised critical state with some maximum number of species per community. In the vicinity of that maximum, a power law distribution of avalanches of local extinction events will occur. As the maximum is approached the net rate of entry of new species slows, then halts.*

3. *On a coevolutionary timescale, coevolving autonomous agents as a community attain a self-organised critical state by tuning landscape structures (ways of making a living) and coupling between landscapes, yielding a global power law distribution of extinction and speciation events and a power law distribution of species lifetime.*

4. *Autonomous agents will evolve such that causally local communities are on a generalised "subcritical-supracritical boundary" exhibiting a generalised self-organised critical average for the sustained expansion of the adjacent possible of the effective phase space of the community*

---

[376] (Kauffman 2000)

[377] (Kauffman 2000) p.160

The first law is really about autonomous agents. The second and the third introduce the edge of chaos distribution and the role of self-organised criticality in keeping the system on the dynamic boundary between static order and chaos. More on them in section 3.1.4.

The fourth law is probably the most innovative of the set. Kauffman rephrases the 4th law in a different way:

> "...*As an average trend, biospheres and the universe create novelty and diversity as fast as they can manage to do so without destroying the accumulated propagating organisation that is the basis and nexus from which further novelty is discovered and incorporated into the propagating organisation*"[378].

The 4th law presents a series of interesting concepts. The first is the adjacent possible. This represents the space of possible innovations that the community of autonomous agents can pursue (consciously or unconsciously) at time t+1, given the set of resources, knowledge and structure of the community at time t. For reasons introduced in various parts of the previous sections, the adjacent possible space is at the same time indefinite and constrained. The exploration of the adjacent possible can be metaphorically visualised as an expanding sphere into the unknown. The sphere contains the set of knowledge, information, routines and practices existing at the multiple levels of the system (individuals, organisations, networks and network of networks) as determined by its history, both at the phylogenetic and ontogenetic levels. The interesting consequence follows that the bigger the stock of knowledge, the bigger the ignorance (intended as space of unknown) with which the community is operationally in contact. According to this metaphoric interpretation, the ignorance becomes the space of possible explorations that the community can perform at a determined time. Because the community is really a set of nodes, relationships and emergent properties, the manners in which the explorations of the adjacent possible is performed depends on the patterns of flows within the community and between the community and its environment. This implies that although in theory the adjacent unknown of two different communities could be similar, the adjacent possible is not. In other words the structure of the community of autonomous agents determines the range of problems the community faces and the range of solutions that the community can achieve at any instant of time.

---

[378] (Kauffman 2000) p.85

The *propagating organisation* constitutes the second aspect worth of attention. The term refers to a self-organising community of autonomous agents, whose organisation depends on the capability of creating further structure of order by dissipating imported energy. As such the *propagating organisation* differs from typical organisations, which survives by adhering to a template specifying the structural and functional relationship among the elements of the system. Instead the self-organising dynamics of the *propagating organisation* (with the inevitable level of redundancy which ensues) makes its survival dependent on the renewal of the processes on which the propagating organisation is based. These are basically patterns of flows along relationship channels. The maintenance of flows and channels at the boundary between static order and chaos can only take place under conditions of exploration of the adjacent possible. If that is not the case, the organisation of the autonomous agents would undergo a radical process of simplification by means of pruning of the least utilised channels and parallel reinforcement of the more useful ones. The propagating organisation survives then by propagation and exploration of diversity. What Kauffman suggests is that the higher complexity of the propagating organisation allows, all other things being equal, a higher dissipative flow of energy by embedding that energy in new structures of order. In other terms, growth allows for higher complexity. At the same time, the capability of absorbing growth depends on the capability of organising the excess variety into a set of flows and processes internal to the propagating organisation. Too fast a change would disrupt the internal balance and precipitate the propagating organisation into chaos. Too slow a change would equally force the propagating organisation toward stasis. Therefore the new meaning of subcritical/supracritical boundary is set in dynamic terms by the maximum amount of exploration in the adjacent possible that the propagating organisation is able to perform without collapsing into chaos.

The difference between the interpretations and use of diversity reported in section 2.5 and in the current one is that in the former case diversity can be broadly synthesised to act as an insurance policy against uncertainty or as a portfolio to promote innovation in the context of a pre-formed system, which may or may not be self-organising, whereas in the latter, the formation of a system comes to depend on the autocatalytic properties of the interactions of a diversified set of initial components and its survival on being on the boundary (and on the rate) of exploration of the adjacent possible.

Chapter 3: Theoretical Framework

3.1 INTRODUCTORY NOTES

This chapter focuses on theory building. Some parts of the content of this chapter have been shown in action, albeit in an anecdotic fashion, in the story of Meadow. I will therefore refer to that chapter, when I feel it useful, to clarify some theoretical points.

The chapter on theory building aims at integrating different literature streams, drawing from complexity theory, literature on industrial clusters, and innovation studies, together with originally developed material in order to a) develop a coherent interpretative framework regarding the origin and development of industrial clusters and b) formulate some empirically testable research propositions. The research propositions do not cover all the theoretical ground but focus instead on extracting from the theory some crucial implications that can be refuted or demonstrated by means of empirical analysis. The final chapter on further research will develop some further research ideas from the theoretical framework.

This chapter is divided into three parts. The first part explores the relationship between self-organisation and power law distribution. I claim that industrial clusters are self-organising systems, and as such, they show a power law distribution in the relevant variables affected by self-organisation. This has never (to the best of my knowledge) been suggested before. The second innovative contribution shows that underlying the three independent models on power law there is the same self-organising dynamic. The discussion suggests the use of power law as an indicator of complex systems dynamic in action. The second part explores the relations between self-organisation and expansion of diversity as a fundamental property of complex systems and clusters in particular. I show that the formation and development of industrial clusters can be explained by the combined effect of four self-sustaining autocatalytic loops. These loops incorporate as basic building blocks several aspects of already existing theories. I also show that each loop contributes in a specific way to the expansion of the overall diversity of the system. Finally in the third part I will offer some reflections and thoughts regarding the demonstrability of the theory. Unless referenced, the theoretical elements presented in this chapter are the author's.

3.2 POWER LAWS, CLUSTERS AND EDGE OF CHAOS

My theoretical and empirical research intends to demonstrate that industrial clusters are self-organising systems. As it was stressed in the literature review this point is both uncommon and controversial: uncommon because the principles of complex systems theory have (somehow mysteriously) not been applied to the discussion of industrial cluster, and controversial because many scholars[379] do not even accept the concept of industrial cluster (as qualitatively different from other forms of geographic industrial agglomerations).

This point raises a series of issues:

1. why do I think that industrial clusters are self-organising systems?
2. why is it of any relevance to show that industrial clusters behave as self-organising systems?
3. how to demonstrate the issues above?

Industrial clusters show a series of properties typical of complex systems. John Casti[380] identified three necessary conditions for complex systems: limited number of agents, local interaction among adaptive agents and agents' decision-making based on (predominantly) local information. A large number of case studies on clusters confirm the three conditions: number of agents (in our case organisations) is between few hundreds and few thousands, interaction takes place predominantly among neighbours (defined geographically and along supply chain lines), and information is mainly acquired in the context of the cluster.

Casti's conditions are necessary but not sufficient. To be truly complex industrial clusters have to exhibit self-organising dynamics and emergent properties. The behaviour of the system should not be reducible to the properties of the single parts, and should be capable of modifying its internal configuration in order to achieve an "*edge of chaos*" dynamic. The evidence provided by several case studies seems to support at least the first points. Coordination for one is an emergent property. For instance, Prato (a textile cluster regarded as the archetype of industrial clusters[381]) is a valid illustration of this

---

[379] (Martin and Sunley 2002)(Amin and Robins 1990) (Belussi 1999) (Whitford 2001)

[380] (Casti 1997) p.231-4

[381] (Piore and Sabel 1984) (Becattini 1990) (Best 1990) (Best 1990; Porter 1990)

point. In 1981 there were 10695 firms in Prato with a total employment of approximately 61.000 people[382]. Although most of these companies were classified as textile, none of them controlled more than 1 or 2 of the numerous production phases of the textile value chain. However, the production cycle requires the vertical and horizontal co-ordination of a bunch of firms, in the absence of a hierarchical or bureaucratic system of control and direction. Coordination of such an extensive network of firms and organisations in absence of hierarchy is clearly an emergent property.

With regard to the second point, any industrial policy aimed at forming, supporting or promoting industrial clusters has to start from the recognition that such systems exhibit emergent properties and therefore require an understanding of the dynamic properties of clusters. But this is not the only consideration: at a more academic level, demonstrating that clusters are self-organising systems sets automatically some constraints upon how the research on clusters should be carried out. For instance, it implies that the use of reductionistic approaches carries the risk of misrepresentation of the system's dynamic.

With regard to the third point, the demonstration of the systemic properties of industrial clusters requires the use of power law theories and some theory building. The demonstration will be done in the following way. I will first make use of the distinction between aggregates and systems and claim that clusters are systems and not aggregates. Second, I will argue that the three power law theories introduced in the literature review represent three aspects of the same self-organising dynamic, as manifested at the three different levels of networks, namely nodes, links and dynamic behaviour of the whole network. Third, I will claim (in line with some literature on power law phenomena) that power law dynamic is an indication of self-organising systems on the "*edge of chaos*". Fourth, in the chapter on results I will show some empirical results concerning power law behaviour in industrial clusters and suggest some further measures to improve and extend the power law discussion in industrial clusters.

3.2.1 POWER LAW THEORIES

In the literature review I have introduced three different power law theories on networks. The first, the *rank-size rule* is concerned with the size and ranking of objects

---

[382] (Becattini 1997) p.535; (Lorenzoni and Ornati 1988) p.45

that in a systems approach can be seen as nodes of a network. The second theory (the most recent), dubbed as the *scale-free network* (SFN), deals explicitly with the topology of linkages among the nodes of a network. The third, the *self-organised criticality* (SOC) theory, focuses on the dynamic behaviour of systems stimulated by endogenous or exogenous perturbation.

All three theories discover the same curious power law behaviour in their unit of analysis. Is this fact a mere coincidence or is it instead an indication that scale-free network, rank-size and self-organised criticality are three aspects of the same thing?

I am going to argue that the three theories are related and that the emergence of a power law can be used first as an indicator of some systemic properties, either at the level of nodes, links or system's dynamic, and, in addition, that those properties are typical of self-organising systems on the *edge of chaos.*

Let me start with an observation: writers on SOC[383] seem to assume that the *rank-size rule* is a manifestation of SOC, without however providing an explanation. Interestingly nowhere in the literature known to the author is this point explicitly discussed[384] (the exception is Krugman[385]).

I'll start with the differences between the three frameworks. First, the three theories start from different units of analysis. The rank-size rule analyses nodal properties (its most celebrated example, the Zipf law correlates size and ranking of cities in the US). SFN describes the topology of network's linkages, and SOC focuses on system-wide dynamic properties. Second, in both the SFN and rank-size rule the description is essentially static and a similar set of assumptions is called for the explanation of the emergence of power law (Krugman's conditions for the rank-size rule and Barabasi's laws for the scale-free networks - see section 2.4.3). SOC instead describes the dynamic patterns of the whole system, that is the reaction of nodes plus links to an exogenous (or endogenous) perturbation. For example, the Saint Andrew fault-line in California defines a particular geographic region where a number of blocks of rocks (our nodes) undergo

---

[383] this is present, for instance, in popular science books, such as (Kauffman 1995; Bak 1997; Buchanan 2000)

[384] A notable example is the (otherwise seminal) book on scale-free network by Barabasi. Although scale-free network theory is on power law, nowhere in the book are the other two power law theories (self-organised criticality and rank-size rule) ever mentioned (Barabasi 2002)

[385] (Krugman 1996)

conditions of dynamic stress and strain (linkages) due to the fault-line's slow movement (which displace the blocks with respect to one another). Stresses and frictions (which play the role of connecting different blocks) link the whole fault-line into a collective system of interaction and feedback (analogous to the critical slope in the sand-pile model). A minor slip between two blocks can induce an amplifying reaction, which could potentially involve the whole fault-line. The resulting effect is bursts of seismic vibrations in which some of the energy accumulated is released. These are plotted on a double log scale and are the subject of the SOC theory (and of course of the Richter-Gutenberg law[386]).

The obvious similarity between the three theories is that they all generate power laws and apply to networks (although networks do not explicitly appear in many examples of SOC) under condition of growth.

Having briefly mentioned the similarities and differences between the three network theories, the way to indicate that they constitute three aspects of the same problem is to find a correspondence between the static structural properties object of the *rank-size rule* and *scale-free networks* theories and the dynamic behavioural properties of the SOC theory.

Let me start from a typical self-organised critical model: the Scheinkman-Woodford model (SWM - see section 2.4.2.1). This model describes a highly simplified supply chain, in which the purchasing and selling of single unit goods take places between agents situated along rows of a grid, representing the different tiers of an idealised supply chain. Firms buy if inventory is empty and sell if goods are in stock. The simulation shows that the market settles to a self-organised critical state, whereby any external fluctuations can trigger a chain of purchases/sales of any length, with a probability determined by a power law. Although it is not explicitly mentioned in the model, I suggest that the length of the chains (or avalanches in the *sand-pile* model) be determined by the spatial distribution of inventory in the supply chain. What do I mean by spatial distribution? A chain of orders branching out from its initial point keeps expanding and propagating until it hits an island of nodes with high inventory. If the distribution of inventory were uniformly scattered across the grid, all chains would have the same length. If instead the inventory distribution were purely random, as in an Erdös

[386] see (Bak 1997)

network[387], then the resulting distribution of chains would exhibit a typical length and be described by a non-power law distribution. It is therefore appealing to suggest that the SOC (dynamic power law) properties of the supply chain system are correlated with a power law distribution of inventory in the system (nodal or structural properties). The SWM shows the parallelism between structural and dynamic properties of power law systems.

An interesting implication of this approach consists in the attempt to redesign (or normalise) networks by transforming dynamic power law (SOC) into static power law (rank-size rule or SFN), by plotting as nodes (or links) the feature that gives rise to the dynamic power law. In the SWM this is the inventory. If an inventory node is defined as the sum of all the adjacent nodes with the same (or very close) inventory level, then the grid would be transformed into a series of adjacent islands of variable size, whose distribution would presumably obey a power law[388]. In this case, the structural and behavioural aspects of the power law would indicate exactly the same thing. The same type of reasoning can be applied to other power law networks. For instance, the same *trick* can be used with the Richter-Gutenberg law. An analysis of the iso-stress lines of a hypothetical fault-line would presumably reveal a power law distribution of iso-stress islands of different sizes[389]. In all cases it is interesting to notice that redefining the nodes as aggregates (along iso-intensity lines) of agents yields a nodal power law distribution. This power law distribution is a direct effect of connectivity forces that shape in a coevolutionary fashion the structural distribution of agents in the network.

Having discussed the equivalence between the three power law frameworks I turn back to the aggregate/system distinction.

It is easy now to see that power law networks are systems and not aggregate. When a network exhibits power law behaviour, this happens because the set of interactions among its constitutive parts generates systemic properties that can not be explained in

---

[387] Paul Erdös was called the travelling mathematician. He formulated a theory of network and demonstrated that connecting randomly the nodes of the network with an arbitrary set of links caused a bell shaped distribution of links. This is characterised by a typical scale with rapidly decreasing frequency around the average value. This is clearly not a power law distribution (Barabasi 2002)

[388] I am not aware that this analysis has been attempted

[389] Still more examples can be given in the field of phase transitions. The distribution of spin if a ferro-magnetic material at the phase transition, as the distribution of solids blocks at the ice-liquid water phase transition, again follows a power law (Barabasi 2002)

terms of single parts' properties. These emergent properties are therefore 'connectionist' properties! Once the agents of an aggregate achieve a sufficient level of connectivity with other agents, suddenly a giant interconnected web emerges. From that point on, the behaviour of the group comes to depend on the web of interconnections and it is no more equal to the sum of the individual behaviours. The coordination achieved in an industrial cluster among independent firms, the spontaneous regulation of the SWM, the collective order established in phase transitions of any types, these are all examples of the emergence of a giant self-organising web of linkages. If the network were an aggregate, no emergent properties could be seen. The previous point provides a useful tool to distinguish between aggregates and systems. If by observing a set of appropriate variables, a power law is found, either in the structural or behavioural properties, then it is possible to infer that we are dealing with a system, not with an aggregate.

The previous considerations on self-organisation, power law, and web of connections lead me to formulate a testable proposition:

Self-organising systems are closer to a power law than aggregates (collections of parts) or hierarchical systems.

Testing the proposition requires some preliminary discussions that will be developed in the methodology chapter (see section 3.3).

3.2.2 POWER LAWS AND CLUSTERS

Industrial clusters have been described in section 2.2. I will just recall few important points. Clusters are geographic concentrations of highly specialised firms in horizontal and vertical relationships of competition and collaboration among them. Clusters grow by an organic process of phase specialisation, which often leads to vertical disintegration and unbundling. A cluster presents the features of highly concentrated and dense supply chains. The units of the dense supply chains are not isolated within their chains, but constitute the agents of a web of economic units that reconfigure themselves into sets of *ad hoc* sub-networks. It is the network, not any single firm, which is collectively responsible for design, manufacturing and commercialisation of products/services. To better describe this approach to organising production and innovation, we can use the metaphor of the organism. Individual parts are specialised, carry out their own functions and try to optimise their fitness. At the same time their interactions in an *orderly whole* build up the organism. So goes the theory or at least part of it[390]. Others[391] are critical of the industrial clusters concept and prefer to see industrial clusters as areas of high concentration of economic activities not dissimilar in their nature from dense supply chains. According to this view, industrial clusters are just one of the forms of the continuum between the Fordist company and the network, and can not be distinguished in principle from dense geographic agglomerations of interactive companies in supply chain relationships. No particular emergent properties are present that can not be explained in terms of dyadic relationships and properties of individual firms. This point raises the issue whether we can use power law to discriminate between self-organising industrial agglomerations, and more traditional agglomerations characterised instead by strong internal linkages (such as, for instance the famous Toyota[392] supply chain). I think we can, at least in a statistical sense, and I will provide a methodology and an example of application.

This methodology starts taking a different unit of analysis, the network instead of the organisation. Seen from the standpoint of the new unit of analysis, the spectrum mentioned above - from organisations to networks – changes into a new spectrum - from aggregates to systems of organisations. The implication is that the new organisational extremes of the network spectrum span from the collection of organisation - random network – to the highly internally organised system of organisations - cluster. As power laws are an indicator of inter-connected behaviour, we can claim that power laws can be used to discriminate between different forms of geographic agglomerations.

This is the idea in synthesis[393]. I take as a unit of analysis the so-called travel-to-work areas in Italy (known as *Sistemi Locali del Lavoro,* in the following SSL). SSLs represent a way to divide the territory into units of self-contained home-work commuting. SLL areas can be classified into 5 distinct types, according to two indexes:

---

[390] (Becattini 1990; Storper 1997)

[391] see footnote 379

[392] Nishiguchi claims that the fractal properties of the *just in time* system in the Toyota supply chain generate self-organising properties (Nishiguchi and Beaudet 1998)

[393] More details are in the methodology chapter (sections 4.1 and 4.2)

industrial specialisation and industrial concentration (both are compared to the national average)[394]. The five groups are defined by increasing thresholds of the two indexes:

- o non-industrialised areas (NIA): no specialisation and extremely low industrial activity (compared to the national average)
- o industrialised areas of type 2 (A2): some specialisation, low industrial activity
- o industrialised areas of type 1 (A1), higher specialisations and higher manufacturing activity.
- o industrialised cluster of type 2 (D2), dominant specialisation of few sectors and high manufacturing activity
- o industrialised cluster of type1 (D1); or so-called super-clusters): extremely high specialisation and high concentration of manufacturing activities.

From NIA to D1 the probability of finding a cluster increases. This does not imply that ID2 and ID1 are automatically to be identified with industrial clusters of the Marshallian type. It merely means that some of the necessary conditions that characterise a cluster are present[395]. According to the earlier discussion on the relationship between power law and self-organising systems, I expect agglomerations which have a higher probability of containing an industrial cluster to show a distribution closer to a power law than other types of industrial agglomerations. In other words, the more the SSL is influenced by the industrial cluster organisational form, the closer the appropriate variables of the SSL are to a power law.

Having specified better the context in which the research will take place, I can complete proposition 1:

Self-organising systems are closer to a power law than aggregates (collections of parts) or hierarchical systems. For industrial agglomerations, the closeness to a power law is related to the degree of self-organisation present in the agglomeration, and emerges in the agglomeration's structural and/or behavioural properties subject to self-organising dynamic.

---

[394] (ISTAT 1997)

[395] This point will be better explained in section 4.1.3

3.2.3 POWER LAW AND EDGE OF CHAOS

In the context of self-organising system, the power law reserves a further surprise. I'll make use again of the Scheinkman Woodford model to illustrate this point. We know that the SWM system tunes itself toward a dynamic state, whereby the length of chains follows a power law. We also know (see section 2.5.1) that the origin of chain lengths distribution lies with the spatial distribution of inventory across the grid. A final point to remember is that a low inventory node allows the propagation of the chain, whereas a high inventory one stops its diffusion. Once the SWM is let free to run, it sets into a state, where the distribution of inventories forms islands (in the hyper simplified model there are only two possibility, node-empty or node-full, let's call them green and red) of varying sizes. In order to get a power law distribution of chain lengths, the size of the islands sizes has to follow a Pareto-like distribution, whereby very many tiny islets are matched by few large 'continents'. This fact leads to an interesting finding: the chaotic and highly fragmented distributions of islets, which causes the diffusion of the chain to propagate in a seemingly chaotic way along green islets, is counterbalanced by the large islands that either (if red) exclude the chain from large part of the territory or (if green) causes the ordered branching out of the chain[396] (see Figure 18, Figure 19 and Figure 20). In the former case the propagation of the chain is chaotic, in the latter is ordered. Order and chaos coexist together at the structural level of the distribution of nodes and at the (correlated) dynamic level of propagation of the chains. Fascinatingly, this is exactly the signature of the "*edge of chaos*" situation, which describes the dynamic state of systems poised at the boundary between order and chaos, where patterns of regularities (which allows structures to emerge) are mixed with uncertainty producing chaos. Furthermore the coexistence of order and chaos is regulated by a power law distribution. We find that the concept of the "*edge of chaos*" is related to a power law distribution[397]. The link between the seemingly uncorrelated concepts of *edge of chaos* and power law distribution suggests the possibility of using a power law as an indicator of a system on the *edge of chaos*.

---

[396] We can metaphorically perceive the emergence of specialisation of territories in this part of the simulation

[397] This result is resonant with Kauffman's approach on Boolean networks. Given a network consisting of a regular lattice of n nodes, and given that each of which has two possible available states, Kauffman discovered that by tuning the average number of connection between nodes, the system shifted from order to chaos, passing through an *edge of chaos* state (Kauffman 1995)

3.2.4 SYNTHESIS

In short, the previous sections have shown the following points:

1. power law distributions emerge in three different types of models; these models can be shown to describe respectively nodal, links and dynamic properties of networks;
2. it is legitimate to think (and I have shown a general reasoning by which this seems to be true) that the three models are really three approaches that point out to the same thing, that is self-organisation in networks. I have shown that dynamic system-wide power law (self-organised criticality) causes a power law distribution in the nodal (or links) properties, suggesting that what is true for one class of models is true for the others;
3. insofar a power law is inherently related to the emergence of a collective system of relationships among the nodes of a system, a power law distribution can be used to distinguish systems from aggregates;
4. this distinction is particularly useful in the discussion on geographic agglomerations of firms. In fact, it is very difficult to assess whether the set of relationships between local firms gives rise to a coherent behaviour. I suggest that the closeness to a power law is an indicator of the emergence of a system from an aggregate of parts;
5. self-organised critical behaviour seems to be strictly correlated with an *edge of chaos* dynamic; the equivalence between self-organised criticality and the other two approaches to network giving rise to power laws permits to generalise the previous point;
6. as industrial clusters are (largely) self-organising agglomerations of firms, the emergence of a power law in their structural properties or in their system-wide dynamic would indicate that they are systems poised on the *edge of chaos*.

In conclusion, the original contributions developed in this section are the following. First, I suggest that the three power law frameworks are related, and that one type of model can be turned into another by a transformation of the appropriate variable. Second, I show that, at least in the case of the SWM, a dynamic power law distribution is correlated with a nodal power law distribution via an *edge of chaos* distribution, thereby suggesting, as Kauffman[398] does, that *edge of chaos* and power law are two sides of the same coin. Third, I develop the proposition that power law distributions can be used as indicators of self-organising dynamic in action, and fourth, I apply that proposition to the specific case of industrial agglomerations and indicate a method to test the above proposition.

3.3 DIVERSITY AND INDUSTRIAL CLUSTERS

The previous section has provided some evidence regarding the correlation between power law distributions and self-organising systems. This implies that a method exists to identify systems characterised by emergent order. I am going to argue in this part that a) diversity at the level of agents is a fundamental pre-requisite for macro-order to emerge; b) the dynamics of diversity is subject to increasing returns, which arise when the various diversity-generating processes link into a giant hypercycle; c) the hypercycle concept allows to give substance to Kauffman's idea of the *propagating organisation* in the context of social systems; d) the *propagating organisation* optimises the product of diversity and connectivity and therefore a measure of connectivity and diversity can be used to test the idea of the propagating organisation.

3.3.1 ECONOMIES OF DIVERSITY

I will in this section expand on the concept of the giant autocatalytic loop of economies of diversity. The first part will decompose the autocatalytic loop in four smaller loops. This part combines concepts drawn from cluster literature, economic geography and complexity theory in a new architecture. In the second part I will discuss how the loops are related to diversity. Finally, the third part will relate diversity and the propagating organisation.

The components of the autocatalytic loop[399] are (see Figure 21):

1. Transaction costs loop
2. Pred/Krugman multiplier loop

---

[398] (Kauffman 2000) p.160

[399] This first part relies on the concepts introduced in part 2.2 (plus sub-sections) of the literature review and in the introduction

3. Shumpeterian competition loop

4. Network reconfiguration loop

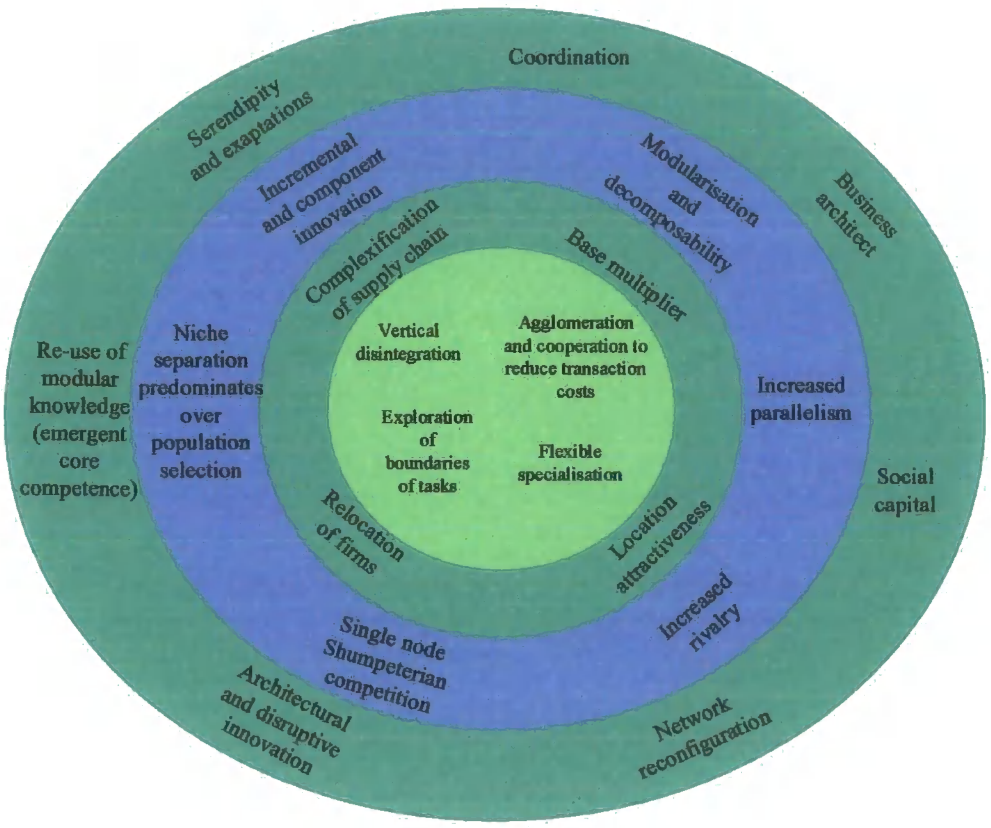


Figure 21: The loops of economies of diversity

3.3.1.1 Transaction cost loop

The history of Hollywood briefly described in section 2.2.7 captures well the meaning of the transaction costs loop. The typical pattern is as follows. A situation of market uncertainty causes a diffused and self-reinforcing vertical disintegration of the dominant organisations (the *Studios* in Hollywood), which in the attempt to reduce risks react by focussing on the most promising and/or safe activities. The outsourcing (or simply shedding of several activities) results in the unbundling of the value chain, which becomes more complex and articulated. The formation of a network of autonomous organisations raises transaction costs. In a cluster this increase is countered by a process of agglomeration that permits the reduction of transaction costs via the creation of an informal community, where information is substantially free and contract formulation and enforcement are largely based on trust. Smaller and progressively more specialised organisations develop a network of linkages through which they cooperate in projects that are often *ad hoc*. This project-based pattern requires specialisation plus flexibility. In their quest for specialisation and agility organisations find convenient to split anytime a possibility to divide a productive task arises[400]. Specialisation-driven fragmentation contributes to the process of vertical disintegration, with the consequence that the overall cycle becomes self-sustaining. From the demand side, the transaction costs loop is an appropriate emergent strategy when demand is volatile, highly diversified, covering overlapping market niches and in general with a low repetitive content. If economies of scale are a critical success factor, a strategy based on the transaction costs loop is not convenient.

In short, uncertainty triggers vertical disintegration and an increase in transaction costs. If this is met with agglomeration and an informal trust-based cooperative business transaction style, this creates the conditions for flexibility and specialisation to set in. This makes further disintegration easier and closes the loop.

The mechanisms described in this section put together the results of the flexible specialisation school (section 1.1) with the transaction cost school (section 2.2.6). The originality consists in proposing that the two analyses are compatible and mutually reinforcing.

3.3.1.2 Base multiplier loop

It is well known in economics that a transaction can have a cascading effect on successive transactions. An initial transaction can feed back into successive tiers of the supply chain creating an aggregated income for the economic system, which could be orders of magnitude bigger than the amount of the initial transaction. If X represents the income of the initial activity, let's suppose that a fraction of it is spent in the supply chain, let's say αX. If the chain is thick and the product is fairly structured, this may lead to a second wave of transaction of which the fraction $\alpha^2 X$ is spent locally, then to a third, a fourth and so on, until all the tiers of the supply chain have been worked out. The sum

---

[400] This exploration of the boundaries of tasks increases flexibility but requires a community where free-riders are easily identified and isolated ((Riolo, Cohen et al. 2001))

of the transactions is the aggregated income Y. The formula for the aggregated income Y[401] is

Equation 1 $Y = X + \alpha X + \alpha^2 X + \alpha^3 X + ... + \alpha^n X$

If the number n of transactions is large enough, this can be approximated with:

Equation 2 $Y \cong X * \dfrac{1}{1-\alpha}$

The multiplier α effectively captures the effects that the initial trigger has on the rest of the local economy as it spreads through its layers. If we call α the fraction transmitted along the supply chain, it is easy to see that the higher α, the higher the size of the aggregated income Y.

If we take a geographically bounded area as our system and the initial transaction as export, we can introduce the idea of base and non-base[402] activities. The former are activities that create value by exporting goods and importing cash, whilst the latter are seen as support to the former. The basic idea behind the multiplier is that the distribution of activities between base and non-base activities can be largely skewed toward the non-base activities, with the coupling between the two provided by α. From a complexity standpoint, the base activities represent the exchange of energy with the external environment that take the system away from equilibrium, and consequently allows the emergence of new structure of order (the non-base activities).

Let's imagine starting from a uniformly distributed economic system without any geographic concentration of the supply chain. Area A receives an order from area B. This base activity triggers a cascade of orders across the distributed supply chain. The hypothesised geographic uniformity of the supply chain represents an unstable equilibrium between competition avoidance – maximising distance from rivals – and transportation costs. If for accidental reasons the volume of the local market grows above a certain threshold, some firms may find convenient to relocate closer to their customers. This move alters the location attractiveness map. In fact, areas with a higher

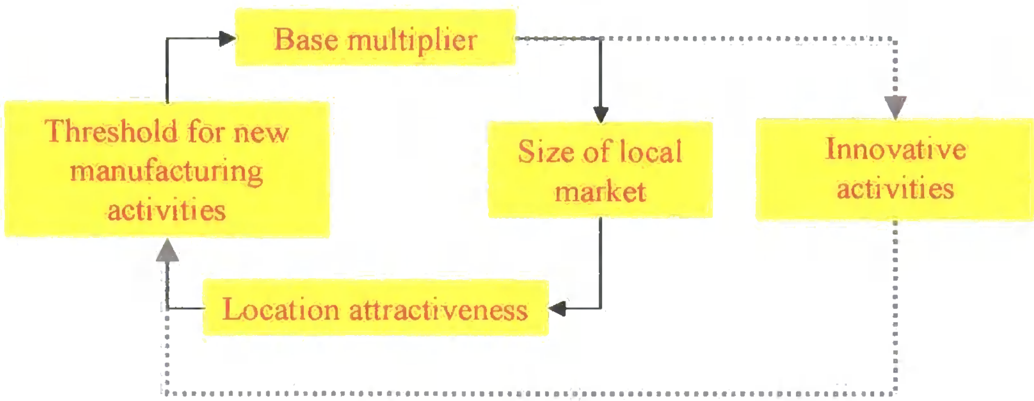
---

[401] This discussion on the base multiplier is borrowed from (Fujita, Venables et al. 1999)

[402] (Krugman 1997)

concentration of firms become more attractive due to reduction in transportation costs and various spillovers effects. The dynamic is circular. The relocation of firms increases local volumes creating further opportunities for more firms to relocate. In turn, the presence of more firms increases the attractiveness of the area.

What happens to the multiplier of the area? I make the simplistic assumption that relocation does not affect the structure of the supply chain. When more firms relocate in the area, the percentage of members of Equation 1 that are part of the area increases[403]. If this fraction is high enough that the approximation of Equation 2 holds, then we can see that, according to the multiplier formula, a concentration of firms in the region increases the aggregated income. The relationship between X, Y and α is approximately linear. In the 60s Pred[404] noticed that the facts were more interesting (and complicated) than the simple description provided by the multiplier formula. When the local economy grows, it becomes more convenient for new activities to relocate locally, as the local economy can support a more diversified set of activities[405]. For instance, as Krugman[406] notices, "*five million square feet is a point of spontaneous combustion. It turns out to be exactly enough to support the building of a luxury hotel. It causes secondary explosions; businesses begin to flock to the location to serve other businesses already there*" On top of that, a further contribution derives from the salaries of the newly recruited workforce, which will be spent locally if the local market offers a sufficient range of goods.



[403] or the leakage of activities outside of the region becomes smaller

[404] (Pred 1966)

[405] a more sophisticated model of urban development based on self-sustaining loops is presented in (Allen 1997) p.45 and 84

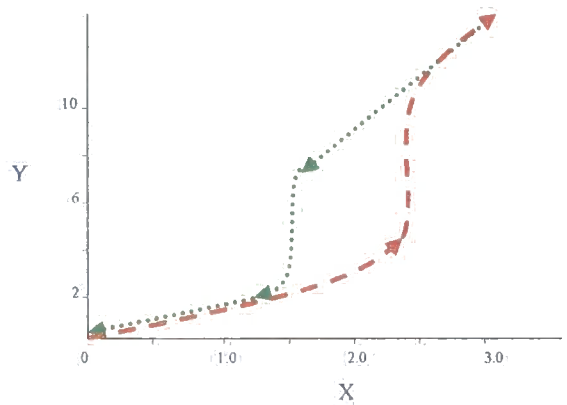
[406] (Garreau 1992) cited in (Krugman 1997) p.244

Figure 22: Pred multiplier of economies of scale

Pred's idea is captured in Figure 22. Pred supposed that the attractiveness function followed a staircase-like behaviour. If for contingent reason a local market overtakes the threshold for new manufacturing activities, then new firms move in, with the effect of triggering a new cycle of base multiplier, which in turn increases the size of the local market and consequently augments location attractiveness. This cycle may be enough to overtake the next threshold and a new cycle will start. The mechanism rests purely on economies of scale.

Key to the ability of the multiplier to deliver its benefit is the concentration of activities behind the final export to be concentrated into a geographic area. The dependence of the multiplier on the size of the local market turns the linearity of the traditional multiplier into a non-linear relationship between multiplier and aggregated income. In fact, the positive correlation between the value of α and the concentration of economic activities within the cluster area indicates that the more the cluster internalises the socio-economic activities that are related to the initial export transaction, the higher will be α and the final aggregated income. The multiplier formula captures the fact that as the size of the cluster economy grows, it becomes more convenient to produce a wider distribution of more sophisticated products and services within the cluster, thus giving rise to an increasing returns dynamics.

Fujita et Al.[407] have inquired into the nature of this non-linearity and have observed a quite complex dynamic around two S curves tied together in a hysteresis cycle (see Figure 23)



[407] (Fujita, Venables et al. 1999), See also (Krugman 1997)

Figure 23: Multiplier's complex behaviour

The figure shows that the expansion of a local economy is initially linear, followed by a very rapid self-sustaining accumulation phase, during which the aggregated income of the region increases dramatically. Interestingly, the onset of the non-linear phase corresponds to a bifurcation in which multiple stable states (green and red curves) correspond to the same set of variables. This is typical of a far-from equilibrium system in which the size of the base activity triggers a self-sustaining dynamic of accumulation. The two curves describe the same qualitative dynamic either (red) in a direction of rapid increase or (green) rapid depletion of Y.

More or less this is the Krugman, Fujita and Venables' account of Pred's multiplier analysis. The nucleus of their explanation lies in increasing returns characteristic of economies of scale at the base of the agglomeration of local economies via the multiplier mechanism. Pred suggested a second loop (see Figure 22), centred on innovative activities stimulated by the agglomeration of firms and activities in the same area, but did not explore the issue and neither does Krugman. I think instead that the innovation loop is central.

We have seen that the size of the multiplier α depends on the size of the local market. In practice this is determined not only by exogenous drivers (location attractiveness), but also by endogenous ones, above all, formation of new firms caused by technological innovations and disintegration of integrated firms in the pursuit of specialisation. These mechanisms, together with Pred multiplier, make the cluster's income dependent on the internal structure of the local supply chains. As a consequence, the aggregated income comes to depend on the mechanisms that make the supply chains thick, complex and internal. I have described in my imaginary story in the introduction how the disintegration of large organisations (into networks of interactive autonomous units) together with location attractiveness and endogenous formation of new firms, caused a net increase in the firm population in Meadow. These mechanisms triggered a self-sustaining vortex of specialisation and Shumpeterian competition, which acted to make the supply chains more complex and fine-grained by multiplying the branches into which the supply chains were structured. The supply chains came to resemble more the delta of a river than a set of linear chains.

The point I want to stress is that the net effect is to strengthen the action of the base multiplier, which, in its turn, by increasing the aggregated income of the cluster, reinforces the conditions for generating further complexity of the supply chains. The base multiplier (in this wider interpretation) captures the fundamental logic of "clusterised" economies, by which seemingly unrelated mechanisms (disintegration, specialisation, location attractivity, ratio between neoclassical and Shumpeterian competition, etc.) interact as an autocatalytic set to promote the survival and expansion of clusters.

The Pred/Krugman's interpretation explains well the fast growth that follows the reaching of the threshold, but remains silent about the reaching of the threshold. The bursts of self-sustaining accumulation that follow the threshold become part of a more complex dynamic. If we consider together exogenous mechanisms (localisation of firms due to increased attractiveness) and endogenous mechanisms (specialisation, innovation and disintegration), then the system acquires an internal dynamic able to propel a self-sustaining process of development. This forms around an internal principle of organisation, provided by the way in which the Pred/Krugman multiplier interacts with the mechanisms that produce the internalisation of activities within the local economy and the increase in complexity of the supply chain. This internal principle of organisation is such that each element reinforces the action of the following one. Disintegration causes Shumpeterian competition, which increases the push for diversification. In turn this makes the chain more diversified, thereby affecting the multiplier. The increase in the aggregated income that ensues generates resources for further innovation and increases the location attractiveness, with the consequence of promoting more competition and forcing a further round of innovation and specialisation.

This giant and complex set of feedback loops will be explored in more details in the next section. What I believe is important is that this set of mechanisms closes onto itself as a giant autocatalytic loop. As all the processes that are part of this autocatalytic loop depend on the internal diversity and connectivity of the local system, the autocatalytic loop itself describes a giant mechanism by which diversity self-maintains and expands.

3.3.1.3 Shumpeterian competition loop

The transaction costs loop together with the base multiplier loop generates an environment populated by small, highly interactive and specialised organisations. In the history of Meadow I have shown how the multiple interdependencies among organisations caused the development of a modular environment, whereby the specialisation (deepening of competencies in a smaller segment of the set of productive processes) was accompanied by the definition of the interfaces between the specialised processes. Modularity leads to increased parallelism, which in turn makes rivalry (between firms at the same level of the supply chain) cut-throat. As the road of a price war[408] is negated by the difficulty of achieving substantial economies of scale (because volatility of market and craftsman type of production make difficult sustaining a long term strategy on price competition), the only way to escape competition is through change and innovation. This type of competition, based on product/process diversification and technological innovation, results in the continuous generation of new niches, through which short-lived competitive advantages (as knowledge on which the innovation is based is easily imitable in industrial clusters[409]) are achieved. If firms attempt the escape rivalry by shifting the terrain of competition towards the creation of new competitive space, this reduces the pressure for the firms occupying the new niches; with the overall result that Shumpeterian competition[410] predominates over cost-based (neoclassical) competition[411]. The generation of new niches is largely due to incremental and component-driven innovation[412] (the other forms of innovations are usually excluded by the small dimension of the organisations and by the limited visibility and control over the sequences of the whole value chain). The overall effect is an expansion of the number of modules, contributing so to close the Shumpeterian loop (see Figure 21).

3.3.1.4 Network reconfiguration loop

The three catalytic loops described above generate a problem of coordination. Disintegration, relocation of new firms and Shumpeterian innovation contribute to an enormous increase in industrial diversity. As the new industrial environment necessitates

---

[408] there is evidence that firms in cluster refrain from price wars for cultural reasons, either because it involves a destructive war with other members of a social community or because it is not considered an option worth of a good entrepreneur

[409] (Rullani 2001) introduces the idea that innovation in industrial clusters is characterised by inclusiveness toward the internal environment and exclusiveness toward the external.

[410] in biology this is known as niche separation effect (see (McKelvey 2003)).

[411] in biology known as population selection effect (see again (McKelvey 2003)). The causes of mortality of firms change: intra-niche selection is reduced but Shumpeterian "creative destruction" due to obsolescence of products/technologies is magnified

[412] for a definition of component innovation see (Henderson and Clark 1990) and (Sanchez and Collins 2001)

the coordination (along the value and supply chain) of a large number of diverse organisations, and as the loops mentioned before do not provide a formal system for coordinating activities, then in absence of hierarchy, it is not clear how the orderly succession and integration of different activities are achieved. There are two answers to this problem. First, proximity provides a frame that encourages repeated interactions (in Axelrod terms: "the window of the future"[413]) and therefore leads to the building of trust. In addition, proximity makes it easier for agents to recognise the members of the community ("tagging") [414], providing an essential condition not only for the formation of trust but also for the punishment of free riding behaviour. In short, history of previous cooperation, decrease in transaction costs (enabled by trust) and increased flexibility (facilitated by the tagging mechanism in establishing new weak-link based relationship), provide a local cognitive map (different for each agent) that helps the development of future interactions. Second, the process of vertical disintegration leaves at the end of the value chains organisations with knowledge of both final markets and production techniques. These organisations, unchained from the ownership of any tangible production asset, become truly knowledge business architects[415] and purport the essential function of interface between the final markets' vagaries and the cluster's organisations. When a new opportunity arises, the business architects organise an *ad hoc* network (a sort of limited life company[416]), which plans and delivers the product throughout the stages of the value chain.

The frequent reconfiguration of the production networks within the cluster represents the organisational response to volatility in demand. In many industries, hi-tech and non, such as electronics, media and fashion, demand frequent and unpredictable changes pressurise organisations to synchronously manage new product development and production. This aspect requires changes in organisational structures to adapt to or accommodate shifting patterns of technologies, products and customers' requirements. Clusters achieve this purpose by keeping the basic units of production (firms) fixed and by changing the aggregation of firms around production and innovation networks. The units (repositories of knowledge concerning specific phases of production) do not change

---

[413] (Axelrod 1984)

[414] (Riolo, Cohen et al. 2001)

[415] In Prato they are known as *Impannatori* (Malone and Laubacher 1998) (Becattini 1997)

[416] (Hall 1999)

frequently (knowledge accumulation would otherwise be difficult), whereas the architecture of the production networks changes according to the competencies, technologies, volumes and degree of innovativeness required by the projects. By acting as the cluster's external structure[417], business architects create an interface between the cluster's organisational practices and the external world's demands and changes. Their role works in two ways: filtering signals in order to anticipate shifts in external demand and selecting relevant information about business and technological practices from which cluster's firms may benefit.

The mechanisms described above are very effective in ensuring innovation and adaptability at the level of the cluster. However, many of the advantages accrue to the cluster and not directly to the individual firms. This fact generates considerable instability. From the point of view of the individual organisation, the shifting pattern of production network represents a considerable investment in new relationships, which involves the need to adapt its organisational and technological structures to new suppliers, customers and partners. Without social capital and the existence of diffused technical tacit knowledge, reconfiguration as an adaptive technique would be slow and expensive.

Reconfiguration has three further effects:

- The emergence of distributed innovation
- The re-use of modular knowledge
- The emergence of exaptation as a mechanism for innovation

Reconfiguration works by recombining modules. Modules' life is usually longer than network's. For instance, specific techniques for weaving woollen fibres may not have changed much in their nature; however the way these techniques interact with complementary techniques regularly changes depending upon the requirements of the specific project. Historians of technology have proposed the analogy with genes[418] and have borrowed the term (technical) *meme*[419] for the role that basic techniques play in the development of products. The frequent recombination of modules generates frequent

[417] see definition of external structure in section 2.3.4

[418] see (Mokyr 1998)

[419] The term meme has been suggested by Dawkins (Dawkins 1989) and indicate a fundamental block of culturally transmissible information/practice.

changes in the interface between modules, with the effect of increasing the probability that two or more modules may discover new synergy regarding the use of existing technologies and/or production processes in new applications. The process is resonant with genetic shuffling. In sexual reproduction, maternal and paternal genes are shuffled to generate new chromosomes that, when recombined in a new individual, generate a different landscape of immune agents, ensuring in this way a higher resistance to parasites. In similar terms, the shuffling of organisations, driven by business architects to satisfy a new project's requirements, creates a series of new interfaces around which solutions emerge at the network level. Innovation is then triggered by the exposure of existing modules to new configurations. As the change affects more the organisation of the modules (architecture) rather than the modules themselves (genes or organisations), the effect is macro more than micro. In complexity terms, it is largely emergent and happening as an effect of the emergence of new interfaces between modules. Therefore this type of innovation affects groups of organisations more than single organisations (the latter case has been treated in the Shumpeterian loop). The combinatorial aspect of architectural innovation explains why innovation at the level of reconfiguration of modules tends to happen at a superior rate than innovation at the level of modules.

Another important element of the catalytic loop is what I have termed recycling or re-use of modular knowledge. Recycling is commonly described as the activity of either reusing an object in a different context or using a degraded form of an object in the same context. The modularisation and high connectivity typical of a cluster creates the conditions by which the same knowledge is re-utilised several times in slightly different contexts (but not too different so as to generate the need for a radical new knowledge to be created)[420]. A specific competence in manufacturing, design, or knowledge regarding the needs of certain market segments is, during the process of forming new production network (reconfiguration), recycled and integrated with other modular competencies. The effect is to generate an emergent macro-competence (truly core[421]), which respects the tests of core competence. In fact, it is practically inimitable, opens new market segments and is based on a complex integration of skills, capabilities and technologies. However, contrarily to the core competencies described in the literature, clusters' core competencies exhibit some original features. First, they are truly emergent as they are

---

[420] See for this point (Maskell 2001)

[421] (Prahalad and Hamel 1990)

formed in the process of reconfiguration of the production networks. Second, they are distributed among independent organisations. Third, the risk of collapsing into "*core rigidities*"[422] is avoided thanks to the temporary nature of the production/innovation networks. As long as the cluster does not evolve into a fixed set of interactive organisations, core competences remain distributed and emergent. The recycling of modularised knowledge therefore gives rise to emergent dynamic capabilities.

The final aspect of the catalytic loop dependent on reconfiguration is exaptation. I have shortly described exaptation in section 2.5.3.3. I will in the following summarily interpret the occurrence of innovation as exaptations as a consequence of the combinatorial nature of the reconfiguration mechanism. Exaptation is based on serendipity. Existing capabilities (in evolutionary biology functions), developed originally as an adaptive response to a set of environmental circumstances, could hide a potential (pre-adaptation) for completely different uses. What I want to stress here is that the serendipitous effect of matching problems to existing solutions depends on a stimulus function. I demonstrate in section 3.3.3 that this function is dependent upon the velocity of recombination of modules in new configurations.

There is a further innovation effect related to the reconfiguration of modules. Innovation at the dominant stage level[423] takes place within the well-defined boundaries of a technological trajectory and as Kuhn would say involves *puzzle-solving science.* Because constraints, environment and customers' perceptions of the technologies/products are well defined, innovation is likely to respond to planning, either within a single firm or in a network of organisations. But, the appearance of unexpected interfaces, which inevitably accompanies the formation of new organisational networks, has the potential of causing via an exaptation mechanism the discovery of new ways of combining existing modules. The effect would be the generation of new dynamic capabilities, that would be architectural in nature.

This effect closes the loop of the reconfiguration autocatalytic set. In fact, the addition of new exaptational functions increases the diversity of the loop and contributes to the coordination problem that got this section started. The cluster reacts to such

---

[422] (Leonard-Barton 1992)

[423] I refer here to the most famous lifecycle model of innovation, that it the Abernathy-Utterback model (Abernathy and Utterback 1978), see also (Teece 1987)

problem by devising new strategies to cope with increasing diversity of modules and functions.

3.3.1.5 Convolution, closure and knowledge

The loops described above do not work in isolation, but are linked in a variety of ways. For instance the modularisation process reinforces network reconfiguration dynamic and vice versa. Architectural innovation by introducing new products tends to redefine the terrain of competition. In fact, the new architecture modifies the context for its components, thereby stimulating new adaptive innovation. What are the overall effect? On the one hand, the autocatalytic loops taken in isolation reinforce the amount and diversity of the loops' components, and, on the other hand, the inter-loops feedback channels create mutualistic conditions of interdependence between different social and economic dynamics. The mutualism accelerates the interdependence between the modules of the single loops and creates a coevolutionary environment, whereby actions do not happen in isolation but have the potential to trigger a series of cascading effects on the interconnected nodes. The resulting effect is the virtual impossibility of isolated change, the transformation of any part without a simultaneous process of adaptation taking place in the rest of the system. Interdependence generates co-evolution that is described by Kauffman[424] as though the evolutionary process was happening on a rubber landscape, where each step taken by any agent modifies the environment of the others.

Coevolution leads to closure. There are several reasons why. Coevolution works by fine-tuning sets of agents' (individual and organisational) responses to other agents' so that the complete set of behaviours resonates as a harmonious ring. Without fine-tuning (and the consequent element of predictability and repetition that ensues) the path of behaviours would swiftly diverge into chaotic and ineffective economic activity. Instead the connectivity and speed of responses between actions provide the circular logic that makes the whole system coherent. Closure is then an indirect effect of the self-referential aspect of coevolution. By closure I don't mean the existence of a barrier to energy or information. Instead I refer to the formation of a boundary condition (referred in the literature review as external structure – section 2.3.4) that filters information and energy flows. In other words, if the parameters against which actions are deemed useful and

---

[424] (Kauffman 1995)

socially acceptable are determined by their effects and if most of those effects are internal, then dynamic proximity will determine the emergence of criteria which drive the conscious and unconscious, individual and collective decision making processes. Closure and coevolution are then aspects of the coherence established across a distributed community. Coherent systems are therefore organisationally close, as norms and criteria (by which decisions are taken and behaviours occur) are defined within the system, but informationally and energy open.

Decisions are based on emergent local criteria, organisational logic is intrinsic, information processing is idiosyncratic, innovation is self-referring; the consequence of all of this is that knowledge is largely tacit, communication is based on local "dialects" with plenty of locally understood words and nuances. The endogenous logic of coevolution and closure then ensures that knowledge, either of productive processes or of social network's engagement rules, is inherently local. This implies that first, knowledge is embedded in a distributed network under the form of organisational routines. These are intrinsically difficult to replicate due to the contextual environmental element in which the integration of skills and theoretical knowledge take place. Second, new knowledge or knowledge about how to do new things is almost predominantly a product of trial and error experimentation, learning by doing and serendipity and therefore very little of it comes in a codified form. This knowledge can eventually leak outside of the cluster's boundary but only in the limit in which it becomes codified. Distributedness and tacitdness are therefore the two barriers to competitive knowledge leakage.

Industrial clusters are distributed systems of innovation and production, which cover a range of technological trajectories. It may be argued that the process of specialisation around phases of productions and the homogeneity of cluster's final products should lead to a process of specialisation around a single technological trajectory. This is instead not the case. Why? The answer is in self-organisation and coevolution. A self-organised network prospers by slowly diffusing away from the techniques and knowledge that contributed to its past success. If migration of capabilities toward innovation would stop, the cluster as a self-organising coevolutionary system would cease to exist and would give rise to an integrated system based on the exploitation of economies of scale. This migration/diffusion occurs by diversifying along two strategies: *shifting* and *deepening* strategy[425]. The former consists in a series of incremental improvements, usually of the component type, within the boundaries of the specific typical production techniques and final products. In the language introduced in section 4.3.1 this is variety increasing. However, reconfiguration of networks may give rise to "better things", that is to new techniques of production, technologies and final products, which open new technological trajectories and are more often based on architectural, disruptive or radical innovation (though very rarely the innovation is ground-breaking[426]). This is known as *shifting* and involves the transfer of resources from old to new uses. Also, under the pressure of Shumpeterian competition the diffusion of capabilities/products from the initial core (as described for instance in the case of Meadow in the introduction) causes the expansion of the range of technologies, techniques and products, so as to cover several related fields connected to the initial specialisation of the cluster. This implies that if the cluster is specialised in, say, woollen textile products, the exploration of complementarities will induce the agents to invest for instance in the completely different sectors of machinery for textile manufacturing, IT and CAD/CAM technologies especially with regard to fashion design, and chemical technologies to make the dyeing process more responsive and agile. The exploration of complementarities with the main specialisation generates multiple technological trajectories. As suggested in a phenomenological fashion in the introduction, the process of building the trajectories (around a distribution of interrelated specific technologies) takes place in a coevolutionary and self-organising context. The absence of macro strategies renders all actions valid in the local context and strongly path-dependent. The exploration of the future possibilities of the trajectories becomes dependent on technological choices carried out to improve individual agents' fitness without any regard for the whole envelope of the technological trajectory. The trajectory itself is self-organising and path dependent. The micro-diversity and connectivity of agents enable multiple paths of parallel experimentations, which in a hierarchical context would be constrained by the footprint established by the macro-strategies.

3.3.2 CATALYTIC LOOPS AND DIVERSITY

I present in this section my view that clusters thrive by exploring diversity.

---

[425] (Ergas 1986)

In the previous section I have described the accumulation dynamics of industrial cluster, achieved by means of reinforcing sets of self-referential processes. One of the effects of the sets is the expansion of diversity. I have represented in Figure 25 the idea that each loop contributes to the expansion of diversity in a specific way.

The transaction cost loop works by unbundling organisations into smaller independent units. In so doing it increases diversity in two ways. First, the process of deepening specialisation increases the organisational focus around a sharper and more defined set of technologies, production techniques and end products. The effect is a modification of the organisations' capabilities. As the process of deepening phase specialisation is achieved by spin-off, the final effect is to reinforce the trend toward an overall increase in the number of organisations. Second, unbundling, by multiplying the number of organisations, multiplies the number of interfaces with which each organisation has to deal with. In fact, unbundling in conjunction with deepening specialisation causes an increase in the number of value chain's phases, thereby pushing up the number of collaborations required to interconnect each phase of production with the adjacent ones. To make an example, if the process of dyeing is split into two parallel sub-phases and each sub-phase is owned by a more specialised organisation, the overall dyeing process will require the coordination of four organisations instead of three, with a consequent increase in interfaces from 2 to 5 (see Figure 24). The two effects work in synchrony to increase what I have defined as interface diversity.

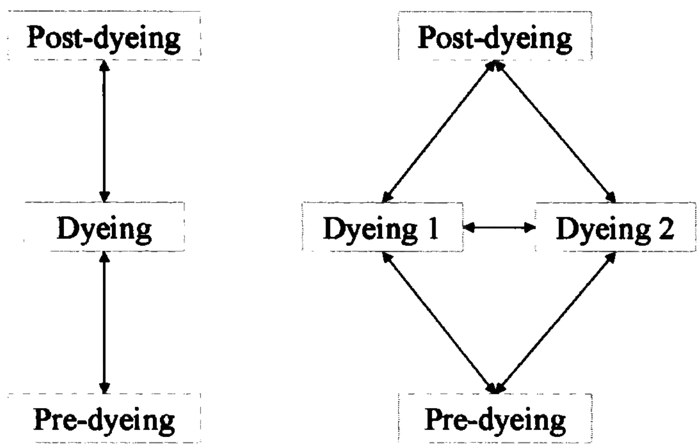


Figure 24: Phase specialisation and interfaces

---

[426] See (Christensen, Raynor et al. 2001)

The base multiplier loop exploits the positive feedback between location attractiveness and market-size driven activities. When the local market hits a threshold, more complex and specialised activities either relocate or start *ex-novo* in the area. As these novel activities are intrinsically different from the existing ones, they contribute to a net qualitatively increase in diversity. The Pred/Krugman base multiplier is ultimately based on economies of scale. Although economies of scale usually tend to increase environmental homogeneity rather than heterogeneity, in this specific case this does not prevent an expansion in diversity as the increase is offset by a reduction of diversity in the nearby environment. This mechanism is different from the transaction cost loop for the following reasons. First, the base multiplier adds new complementary stages to the existing phase of the value chains, rather than fractioning the existing ones. Second, the diversity increase affects the dimension of disparity more than variety and balance (see section 4.3.1). Third, the process is not purely endogenous.

The case of the Shumpeterian competition loop contribution to diversity is again different from the previous two. The increase in diversity involves the creation of new technologies/products and consequently new organisations. Whereas the first mechanism may involve some process innovation, which is reflected in the spinning-off of new organisations, and the second concerns the relocation of start-ups or of new organisations around existing technologies/products, the Shumpeterian case involves process and product innovation. Organisations escape the tyranny of neoclassical competition by moving away from the terrain of competition itself (see section 3.3.1.3). As competition takes place between individual organisations at a similar level on the value chain, innovation will happen at the level of single value chain activities, that is, at the modular level. The effect can be process or product innovation. If innovation is purely substitutive, then it does not contributes to total diversity. Alternatively, innovation adds to existing diversity by expanding the range of intermediate and/or final products/processes. The contribution to diversity is mixed. As most of the innovation is modular and incremental, this mechanism adds more to variety than to disparity.

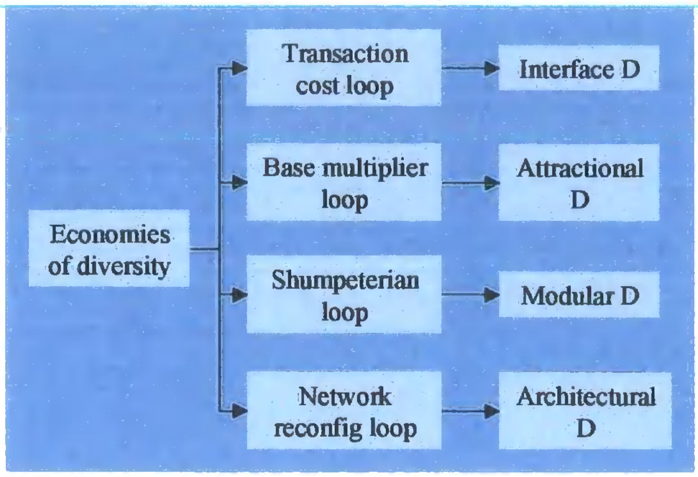


Figure 25: Autocatalytic loops and the expansion of diversity

The fourth autocatalytic loop generates diversity mainly by architectural and disruptive innovation. Whereas the Shumpeterian competition loop innovation regards the single node of the network, the network reconfiguration loop acts more at the network level. From this point of view, this type of innovation constitutes an emergent property of the dynamic of forming and disbanding of production networks. The contribution of this loop to diversity is multifold. First, the coordination problem (introduced in section 3.3.1.4) necessitates forms of emergent integration of networks that we have defined as business architects. Business architects are organisational forms that introduce a level of hierarchy in the cluster. By regulating part of access to final markets and by integrating modules into production networks, business architects are more akin to regulatory genes than to Mendelian genes[427]. Clearly the emergence of business architects represents a fundamental step in the evolution of diversity. In fact, all previous loops add to the diversity of organisation. Business architects contribute instead to the "*organisation of diversity*"[428] and represent a mechanism of auto-regulation of diversity itself. Second, architectural innovation has a more radical character than modular innovation. As such it tends to open new (either product or technological) trajectories, thereby adding more to disparity than to variety.

[427] Mendelian genes are responsible for the expression of a single phenotypic trait (for instance blue/black eyes). Regulatory genes instead control networks of genes, which thanks to a complicated mechanism of inhibition and activation, control complex metabolic pathways.

[428] The expression is by Grabher (Grabher and Stark 1997) p.536

We can roughly divide the contribution to diversity of the four previous catalytic loops into two camps: the first two add to diversity via organisational changes, the second two via the introduction of new technologies/innovations.

3.3.3 THE SPACE OF EXAPTATIONS

I have introduced the concept of exaptation in the literature review and briefly reviewed it in the discussion on the network reconfiguration loop. I think that industrial clusters constitute a preferential space for exaptations to happen. Let me explain why.

The mechanism of exaptations depends on the supply of: a) random mutations, neither harmful nor beneficial to the bearer, and able to persist in time and spread in the population; and b) adaptations, that is mutations that make the bearer's chance of survival higher. Both are likely to turn into exaptation when (usually in response to an environmental change) the mutation or the adaptation finds a new life in a different application or function.

Exaptations and diversity are connected at least under two aspects, the supply of mutations and the transformation rate of mutations into innovations (from now on called actualisation).

With regard to mutations (which in history of technology can be equated to invention[429]), if we make the oversimplifying hypothesis that the actualisation rate is simply dependent on the amount of available mutations, the question becomes whether there are conditions under which the supply of mutations can be above average and self-sustaining. The rate of production of inventions will in general depend on a) the amount of directed micro-exploration performed by agents, b) the amount of non-directed micro-exploration (the probability of stumbling upon an unexpected phenomenon whilst performing other activities), c) the number of agents and d) the speed of diffusion of information among the agents[430]. These terms can be expressed as:

Equation 3 $$M = E*F*N_A$$

---

[429] (Mokyr 1990)

[430] clearly other factors are important, such as, quality of education, nature of competencies, entrepreneurial spirit, etc.

E is the rate of micro-exploration (directed and non directed) performed by agents, N_A is the number of agents and *F* represents the information flux. Clearly the terms in the equation are fairly complex. For instance the factor E includes all sort of path-dependent parameters, including the rate of success of past inventions, the perceived reward from innovators, specific elements of science and technology, etc. The factor F instead describes the connectivity of the agents, the openness of agents to diffuse knowledge, the frequency and density of social transactions, etc.

The capacity to innovate due to exaptations will then be a function of the transformation rate of mutations into innovations. We can express it as:

Equation 4 $$\mathbf{I = M*A}$$

I stands for innovation capability due to exaptations, *M* is the rate of generation of exaptations, *A* represents the rate of actualisation of exaptations. The basic idea behind the actualisation of exaptation is that this is likely to happen when a solution evolved in a specific context becomes part of a different context. To make my reasoning clear I will define some terminology. A context is a specific micro-environment (biological or industrial or technological)[431] where adaptations/mutations (functions or capabilities or techniques) evolve endogenously or in response to environmental pressures. This micro-environment is made of a set of modules (species or industrial agents or technologies) with which the adaptation/mutation is interacting either through direct (strong links) or indirect coupling (weak links). The environment at time t+1 is the path-dependent result of the interactions of adaptations/mutations and modules in the micro-environment (context at time t).

The actualisation factor A is expressed as the product of the micro-actualisation rate $\underline{a_{I\sigma}}$, which expresses the probability that an adaptation/mutation $\underline{i}$ will give rise to a new function $\underline{I}$ in the process of interacting with the module $\underline{\sigma}$. The probability of such interaction is given by $\underline{c_{i\sigma}}$ which expresses the probability that function $\underline{i}$ will interact with module $\underline{\sigma}$. Summing over all the modules $\underline{\sigma}$ (the sum of the contexts gives the specific micro-environment of function $\underline{i}$), we get the rate of actualisation for the exaptation $\underline{I}$. Finally, taking the sum over all exaptations, we get the total factor A.

---

[431] the micro-environment coincides with Maturana's concept of external structure (see section 2.3.4)

Equation 5[432] $A_I = \Sigma_\sigma a_{i\sigma} * c_{i\sigma}$

Equation 6 $A = \Sigma_i \Sigma_\sigma a_{i\sigma} * c_{i\sigma}$

To make an example, the exaptation flight seems to derive from the adaptations feathers + wings (which existed before flying evolved). If we refer to the exaptation flight as I, and to the adaptation feathers + wings as i, then flight (I) emerged when an agent carrying i, was exposed to a different context σ, probably free fall (for instance a prey forced to jump from a tree threatened by the sudden appearance of a predator)[433]. The pre-existence of feathers and wings was discovered to be useful to survive in the new tree-free fall-predator context. If we express the context as a modification of the agent's external structure, we see that the emergence of the exaptation is related to the exposure of the capability i to the sum of modules σ (predator, free fall, etc.).

I suggest that at least part of the high innovative capability that industrial clusters show is due to exaptations. In fact, if mechanisms for increasing returns in the supply of technology exist, then, these could explain why innovation activities seem to concentrate in few, sometimes very small, geographic spots around the world (see section 2.2.7). This phenomenon is not restricted to technological innovation only. History of art provides some striking examples. The incredible flourishing of artistic activity in Renaissance Florence or in classic Athens is a testament to the power of localities, not less than the stunning story of today's Silicon Valley.

Why does agglomeration lead to innovation? Let me go back to the two mechanisms I have mentioned to explain the occurrence of exaptations: production of mutations/adaptations (in history of technology the role of mutation is taken by any change in the stock of knowledge or techniques, and adaptations by existing capabilities) and actualisation (conversions of mutations/adaptations into new functions).

---

[432] the equation takes into account only first order interactions. It is plausible that conversion of invention in innovation takes place when two or more modules (capabilities) interact with the exaptation. In the case of two interactions the equation becomes

$A_I = \Sigma_{\sigma k} a_{i\sigma} * c_{i\sigma} * c_{ik}$

[433] The alternative hypothesis regarding the origin of flight is that wings helped a potential prey to run faster uphill

I will use the equations I have described earlier to make my point. The first term in Equation 3 – exploration – is dependent on several factors, such as the nature of technology, the stage of technology lifecycle, the intensity and type of competition, just to name a few. Focussing on the competition aspect and using a biological analogy, species competing for resource (either in a commensalistic or prey-predator fashion[434]) can either drive each other out (population selection dynamics; in economic terms price competition) or, by a process of specialisation, create new niches, thereby changing the terrain of competition. The latter process gives rise to Shumpeterian competition, by which new technological species are generated and a more complex environment evolves. Generally, if Shumpeterian innovation generates only short term competitive advantages, as it is the case when a high number of actors have the capability to understand and imitate the innovation, competitive pressure will not subsume, and the system will keep innovating in order to escape the ever mounting population selection pressure. The dance between the two types of competition under condition of high diffusibility of imitability of knowledge gives rise to increasing returns in the supply of the first term of Equation 3.

The other two terms in Equation 3, the flux of information and the number of agents, show above a certain threshold a self-reinforcing dynamics. In fact, fast diffusion of information causes learning and imitation. This in turn triggers a competitive innovative dynamic, thereby creating an environment that can support a higher population. This description is similar to the one highlighted in the Shumpeterian loop.

If the Shumpeterian loop is the main source of mutations/adaptations, the network reconfiguration loop is the main cause of actualisation of mutations/adaptations. Let us see how. In general, the higher the speed of generation of new interfaces between modules, the higher the probability of the occurrence of actualisation. Actualisation is fundamentally a combinatorial game. Hence an environment internally structured in multiple interactive niches will be more conducive to actualisation than a homogenous one. The reasons are the following. First, the ability to do something depends often on the ability to do a spectrum of other things; second, the ability to learn depends on previous learning; third, the ability to invent depends on the capability of establishing connections among apparently non-related things. Therefore, an environment where a distribution of connectivity channels allows the newly formed mutations to interact with a

434 (Maynard Smith 1974) (Saviotti 1996)

rich variety of modules (or capabilities) increases the chance of conversion of inventions/capabilities into innovations. This means matching supply of inventions/capabilities with needs that may exist at the pre-conscious level[435] (as it is often the case with radical or disruptive innovation). These 'invisible' or 'dormant' needs escape the logic of market-pull and marketing planning and necessitate a phase of exposure to the novelty through which, most often by pure chance, a critical connection is made with regard to the use of a new or existing technology/idea in a new area. The network reconfiguration loop describes an environment in which frequent reshuffling of modules in new configurations generates new microenvironments for the agents' capabilities. The capabilities, products and mental frameworks of existing modules happen to interact with different modules in a new configuration. As reconfiguration creates new interfaces between modules, adaptations have a higher chance to become exaptations and "dormant" mutations can become source of innovation and competitive advantage[436].

So to conclude this section what have diversity and connectivity to do with exaptations? Well simply:

1. exaptation constitutes one of the sources of diversity (at least on the supply side), it provides a regenerating platform of mutations/inventions that feed into existing practices, routines and technologies;
2. under certain conditions, the production of mutations/adaptations follows a non linear, self-reinforcing, dynamics, which I have identified and described as the Shumpeterian loop (see section 3.3.1.3);
3. environmental diversity, that is the degree of structuration of the environment into overlapping niches, provides a fertile setting for the '*eureka*' phenomenon to happen; the more diverse the setting, the higher the probability of actualisation;
4. the speed and density of social transactions (communication of information/knowledge and exchange of tangible goods[437]) facilitates the matching

[435] needs that can not be quantified in terms of economic demand

[436] In the introduction I told the story concerning yoghurt as a piece of technology/knowledge used for household consumption, that when introduced in the thriving Meadow's environment was transformed into a new industrial product. This is an example of a transformation of a "dormant" capability into an exaptation

[437] (Boisot 1998)

of supply of mutations with demand of solutions. Connectivity is therefore crucial to the formation of a spiral where diversity, connectivity and exaptations work together in a self-reinforcing way.

3.4 HOW TO DEMONSTRATE THE THEORETICAL FRAMEWORKS?

The purpose of this section is to extract some propositions to be tested in the context of industrial clusters. In the following I will recap the main points of this chapter in order to show the emergence of the propositions from the theory. These points are condensed and graphically presented in Figure 26.

1. The initial point is really made by Kauffman. "... *diversity probably begets diversity; hence diversity may help beget growth*"[438]. Kauffman suggested, using computer simulations, that, given a set of components and some rules of interactions, diversity of initial components has a finite probability of leading to the synthesis of further components, ending up in further diversity. There is probably a threshold above which the expansion of diversity becomes self-sustaining. The examples, given in the literature review (section 2.5.2) showing the self-reinforcing dynamic in the development of agriculture from the hunters-gatherers society, constitute valid evidence of the mechanism.

[438] (Kauffman 1995) p.292

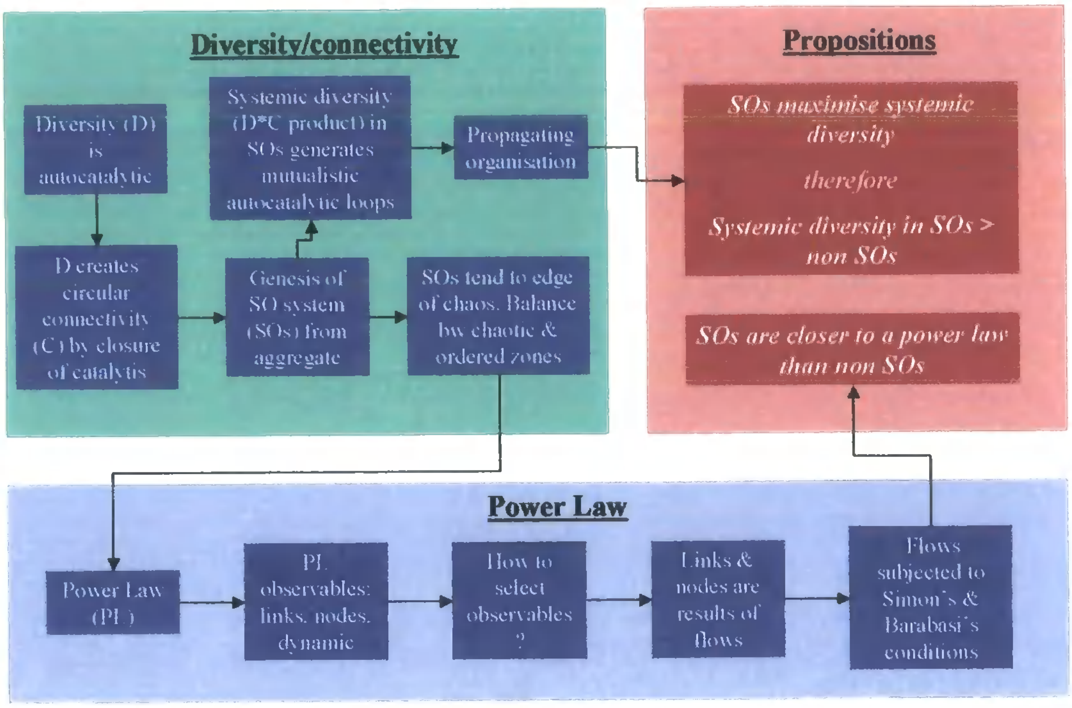


Figure 26: Main theoretical points and propositions

2. The expansion of diversity is likely to result in the closure of some catalytic cycles. The closure causes a series of effects: a) it determines the emergence of an internal architecture of connectivity based on dynamic self-sustaining processes, b) it provides a preferential path of development by providing rules for selection and exclusion of processes, c) it introduces an internal dimension and dynamic, thereby giving rise to the emergence of a unique identity. The closure introduces internal rules of organisation. The emergence of the rules causes the transition from an aggregate of parts into a system, whose parts are subjected to a coevolutive dynamic. Coevolution and closure are conjugated variables. Closure and coevolution are structured around some internally determined rules of organisation, which act to decouple the system from its environment. The system becomes information and energy open but organisationally closed. This means that the rules of cognition and engagement with the environment are fixed by the internal dynamic. The modification of the rules happens as a response to internal reorganisation, which may or may not be triggered by external events. In other terms the organisational rules are genotypic, whereas the space of implementation of the rules is phenotypic. The distinction between internal dynamic and external interactions generates a filter between the environment and the

system, something that decouples the external from the internal dynamic, and allow the appearance of self-organising properties.

3. Self-organising systems mix chaotic and ordered features as if both types coexisted in one. Chaotic dynamic allows for frequent reconfiguration and emergence of novelties, ordered dynamic allows for robustness and homeostasis. The balance between the two is described by the metaphorical expression of *edge of chaos*. At the *edge of chaos* the distribution of connectivity alternates strong links within highly connected sub-networks with weak links connecting the various local networks (within the system).

4. The distribution of links within and across sub-networks and the dimensions of the highly connected sub-networks obey a power law. The emergence of a power law is an indicator of self-organising dynamic. In fact, a power law indicates that the system tunes itself toward a state whereas the distribution of structural and behavioural properties is such that all scales and behaviours of relevant network's properties are present in the system.

5. Power laws emerge in the three constituent elements of networks, i.e. links, nodes and behaviour. Therefore a research that intends to test the presence of power laws should first select the appropriate variable. I have discussed previously the fact that, in my opinion, the appearance of power laws at the level of the three observables (links, nodes and network's behaviour) constitutes evidence of the same underlying dynamic, that is, the self-organising nature of networks at the *edge of chaos*.

6. Still the selection of the appropriate unit of analysis is complicated by the fact that the recognition of power law behaviour has occurred independently from one another and in the context of three different theoretical frameworks – rank-size rule in economic geography in the 30s, self-organised criticality in physics at the beginning of the 90s (readily embraced by complexity scientists) and, scale-free networks (again in physics) at the end of the 90s. I have suggested that a transformation of variables can change for instance a nodal variable in a link variable, thereby raising the possibility of using indifferently any variable.

7. Here is the main line of reasoning: diversity is instrumental to the formation of self-organising systems; in such systems the balance between ordered and chaotic features follows a power law; the three classes of power law theories describe the three

fundamental aspects of networks. In short, self-organisation is explainable in terms of the set of interdependencies that determine the set of flows between nodes and links. This is something that a statistical analysis on appropriate systems can test. If self-organisation and power law are correlated then we can use a power law to test (or reveal) self-organisation in action. However, a dichotomy between self and non self-organising systems is difficult to find. Therefore our proposition will have to test a difference in degree more than kind. From this point of view, a comparison between systems characterised by different degree of self-organisation should reveal a correlation between self-organisation and closeness to power law.

Proposition 1. Self-organising systems are closer to a power law than aggregates (collections of parts) or hierarchical systems. For industrial agglomerations, the closeness to a power law is related to the degree of self-organisation present in the agglomeration, and emerges in the agglomeration's structural and/or behavioural properties subject to self-organising dynamic.

8. The proposition above tests the relationship between power law and self-organisation in networks. It is however silent with regard to diversity as a morphogenetic feature of self-organising systems. Eigen and Schuster showed that the closure of catalytic reactions formed the origin of systems. Kauffman showed that diversity is critical in achieving closure. However, the contribution of diversity does not stop at the genesis of systems. I have suggested in this chapter that clusters (as example of self-organising systems) can be explained by means of a set of interdependent autocatalytic loops. Each loop finds its roots in the interaction of diverse parts and stimulates the generation of further diversity. Systemic diversity is different from diversity prevailing in aggregates. In fact, the former is the result of the expansion of the system in the *adjacent possible*, constrained by path-dependency and occurring along a number of technological trajectories, the latter is merely fortuitous. In other words the expansion of diversity in systems is a result of entangled processes. The entanglement is an effect of the connectivity of the system and expresses the idea that the actions of the agents can not be separated from the other agents' actions and from the environment. Briefly, systemic diversity and connectivity are inextricably linked, whereas this is not the case with aggregates.

9. A self-organising system and specifically an industrial cluster is therefore characterised by the following properties. First, it is dissipative, that is the system can not be isolated from the set of industrial, informational, social and economic gradients with which is connected. Second, the dynamic of the system is based on the set of mutualistic autocatalytic cycles discussed above. Third, the properties of agents as nodes of the network and their relationships as links follow a power law distribution. Fourth, the structure of the system is not based on a static architecture independent from the agents' relationships. Instead, architecture and processes coincide[439]. As the architecture is constantly rebuilt and revitalised by the agents' exploration processes, it starts to emerge that this architecture, and really the organisation itself, consists of the multiple processes of exploration that take place in the context of the reconfiguration of modules. It is the product of diversity and connectivity. Seen from this angle, Kauffman's rather indistinct definition of the *propagating organisation:* "*...As an average trend, biospheres and the universe create novelty and diversity as fast as they can manage to do so without destroying the accumulated propagating organisation that is the basis and nexus from which further novelty is discovered and incorporated into the propagating organisation*"[440], becomes more precise and subject to empirical analysis. The features of the *propagating organisation* are the four ones (dissipativity, mutualistic autocatalytic cycle, power law distribution and dynamic architecture of reconfiguration of modules) reported above. The *propagating organisation* is at any instant of time the product of the processes of exploration times the connectivity that makes those processes possible. Considering connectivity together with diversity prevents the common mistake of isolating the diversity component and looking for the causes of changes in diversity[441] alone. Stated another way, my research strand on the properties of self-organising systems has converged with Kauffman's research for a fourth law of thermodynamics, one that takes into account the properties of the living to co-create the environment in which they happen to be. Self-organising

---

[439] (Lane 2001)

[440] (Kauffman 2000) p.85

[441] usually intended as cultural, ethnic, gender, etc. The maximisation of diversity is described as to leading to creativity, see (Thomas and Robin 1996) (Leonard and Strauss 1997). A further example of diversity being isolated from connectivity, that is the discussion in the economic geography between economies of localisation and economies of urbanisation, is presented in the conclusions (section 7.2.1).

systems maximise the product of processes of exploration at the agents' level with changes in the connectivity in which (and by means of which) the exploration can take place. In terms that makes sense in the context of my research, this statement can be reformulated as:

Proposition 2. Self-organising systems maximise systemic diversity (product of diversity times connectivity) at a rate higher than aggregates (and hierarchical systems).

Chapter 4: Methodology

4.1 THE UNIT OF ANALYSIS

4.1.1 THE CONTEXT

The purpose of this chapter is threefold: first, to illustrate the methods of research adopted to test the propositions introduced in the previous chapter (theory), second, to provide a rationale for those methods and third, to expose in some detail the procedures and steps taken to validate the propositions.

For clarity the propositions are given below.

Proposition 1. Self-organising systems are closer to a power law than aggregates (collections of parts) or hierarchical systems. For industrial agglomerations, the closeness to a power law is related to the degree of self-organisation present in the agglomeration, and emerges in the agglomeration's structural and/or behavioural properties subject to self-organising dynamic.

Proposition 2. Self-organising systems maximise systemic diversity (product of diversity times connectivity) at a rate higher than aggregates (and hierarchical systems).

Proving (or disproving) the propositions requires first the identification of an appropriate unit of analysis, that is, a terrain of analysis with the following characteristics:

1. distribution of discrete social and industrial agglomerations;
2. distribution of agglomerations ranging from aggregates to self-organising systems;
3. availability of populations large enough to allow statistical analysis, sampling and significance tests.
4. availability of reliable data

The first condition demands that the unit of analysis includes geographically circumscribed industrial agglomerations. The word discrete implies the possibility of drawing boundaries. The second condition demands that the unit of analysis covers a range of organisational forms that range from the self-organising industrial clusters[442] to simple aggregates of firms. It is very plausible though that each type of agglomeration presents a different mix of self-organisation and hierarchy, or in other terms of properties endogenously evolved in a bottom-up fashion, and properties dependent upon the imposition of a monolithic network structure. The third and fourth conditions are self-explanatory.

The choice of the unit of analysis fell upon the so-called travel-to-work areas in Italy (known as "*Sistemi Locali del Lavoro*" or SLL).

Why Italy and why travel-to-work areas? The choice of Italy is dictated in part by reasons of convenience (the author is Italian and has been able to put to use a network that has granted him access to data and knowledge otherwise more difficult to obtain), and in part by the specific importance that Italy has represented in the discussion on industrial clusters as an alternative form (post-Fordist) of industrial organisation (see literature review, section 2.2). The second aspect, the selection of travel-to-work areas, requires some additional notions with regard to what travel-to-work areas are and why they are an appropriate units of analysis.

4.1.2 I *SISTEMI LOCALI DEL LAVORO* (SLL) OR TRAVEL-TO-WORK AREAS

Travel-to-work areas (SLLs) represent a subdivision of a territory in self-contained areas of home to work commuting. The idea[443] and related statistical algorithms were developed in the 70s by a group of statisticians, economists and geographers working at the IRPET[444] and at the University of Newcastle upon Tyne in Britain. These techniques permit the calculation from Census data of the geographical distribution of localised economic activities at a fairly detailed and granular level. The assumption behind the SLL idea is that the geography of home-work commuting is related to the social and economic environment in which commuting takes place. The basic idea is that the higher the percentage of internal home-work commuting taking place within the boundaries of an area, the higher the chance of capturing within the area some territorially-specific

[442] In the chapter on theory I define clusters as self-organising systems

[443] SLL are known in Europe under other names: in Spain, *mercatos locales del trabajo*; in Britain, *local labour markets*; in France, *zones d'emploi* ((Brusco and Paba 1997) p.272)

[444] Istituto Regionale per la Programmazione Economica della Toscana – Regional Institute for the economic programming of Tuscany

social and industrial aspects. The SLL represents an algorithmic way to define the micro-unit of analysis of economic geography and economic sociology.

4.1.2.1 Some methodological issues

Traditional economic models[445] do not consider geography as an important variable in the characterisation of economic activities. However, constraints to mobility, transportation costs, the stickiness of knowledge and industrial practices (all terms we can collapse under the heading of local histories) create a multiplicity of local economic and industrial cultures that are reflected in geographic patterns of specialisation. Although the observation that localities show idiosyncratic aspects is common sense, finding some operational criteria to divide a macro-territory into discrete units (which maximise the expression of those specifities and allow the geographic study of those units) is exceedingly difficult. In general the study of local production systems can be approached either in a qualitative or in a quantitative way. Most of the studies in the literature[446] adopt the qualitative case study approach. It can be argued that this is a powerful method for hypotheses generation, but it is unsatisfactory for hypotheses testing. In fact, apart from the problem of validation of case study results, the non-comparative nature of case study methodology and the non-random nature of sample selection, make difficult the empirical verification of general hypotheses[447]. The alternative approach relies on the use of statistical and econometric techniques. There are two main directions here[448]. First, some studies try to define agglomerations by studying patterns of diffusion of innovations in order to define the geography of knowledge spillovers. The alternative approach examines economic output measures, such as employment patterns, productivity, rate of growth, entrepreneurship, etc[449]. It tries to determine whether output measures show a geographic dimension. The assumption is that agglomeration dynamics create a symmetry break in the homogeneity of distribution of economic activities. The SLL approach falls in the latter category.

---

[445] (Fujita, Venables et al. 1999)

[446] the famous book by Saxenian (Saxenian 1994) on Silicon Valley is a paradigmatic example of this approach

[447] (Engelstof, Jensen-Butles et al. 2002) p.15

[448] (Almeida and Kogut 1997; Feldman 1999)

[449] (Glaeser, Kallal et al. 1992), (Henderson 1986)

4.1.3 CLASSIFICATIONS OF INDUSTRIAL CLUSTERS

4.1.3.1 Sforzi-Istat classification

In 1981 the implementation of the SLL approach divided the national territory in 955 areas. Ten years later the SLLs were considerably less: 784[450]. When the first data became available, a group of researchers lead by Fabio Sforzi started asking some quantitative questions about industrial districts: how many there were, where were they, what was their combined value in GDP terms, what were the patterns of employment, etc. The econometric and statistical analyses provided by the group, together with the support of the national institute of statistics (ISTAT)[451], managed to put on a stronger base the results achieved by the various schools studying industrial clusters (see literature review, section 2.2) using more qualitative approaches. In fact, since then the work has become the reference standard for econometric and geographic work on industrial cluster[452]. This work is known as Sforzi/ISTAT classification[453]. Sforzi started from the notion of the Marshallian district and extracted from it some necessary conditions that capture some elements of the dynamic underlying the interactions between small firms in a district. Essentially these have to do with the phase specialisation mechanism (described in the introduction and in sections 2.2 and sub-sections), which generates and perpetuates a high concentration of small firms, highly specialised and interactive. In detail the criteria that Sforzi'[454] identifies are: a) percentage of workers in the manufacturing sector, b) percentage of workers in manufacturing SMEs and c) percentage of workers in those sectors in which the SLL results specialised. Given (somewhat arbitrarily) some thresholds for these criteria, SLLs can be dichotomically divided into two classes, clusters and non-clusters. For each cluster a main sector of specialisation is indicated. The procedure is internally consistent although it contains some arbitrary elements especially in the determination of the threshold levels. It is important to stress that the classification of an SLL as a cluster only indicates the presence of some necessary conditions for the existence of a Marshallian district. It does

---

450 I sistemi locali del lavoro 1991 (ISTAT 1997)

451 Istituto Nazionale di Statistica (http://www.istat.it)

452 (Cannari and Signorini 2000) p.3

453 I sistemi locali del lavoro 1991(ISTAT 1997)

454 (Brusco and Paba 1997)

not identify for certain the presence of an industrial cluster, but merely indicates that the SLL's social and productive system is compatible with the nature of a cluster[455]. In other words, the algorithm does not identify a cluster, but lays down a way to identify those SLLs that may harbour a cluster. The identification of clusters requires research technique more sensitive to the complexity of local situations. The great contribution of the Sforzi's work is that it provides a base for quantitative comparative analyses of territorially bounded industrial agglomerations.

This approach is open to multiple sources of errors. The first and simplest is quantitative. Trying to capture complex realities by using common thresholds is equivalent to imposing a straightjacket to situations that are by definition idiosyncratic[456]. Second and more serious, the procedure of identification of SLLS and the dichotomic Sforzi/Istat can only capture situations in which there is a clear dominant specialised sector. If the sector of specialisation is covered by the "noise" of larger sectors, then the cluster could remain unnoticed[457]. Also, by focussing on traditional SIC codes, the procedure can not take into account the fact that clusters are cross-sectors. Thirdly, if the percentage of employees involved in cluster's sectors is small compared to others, the SLL procedure may not be sensitive enough to locate clusters spread over several SLLs. The Sforzi/Istat calculation procedure therefore is like a low-resolution camera, which transforms a complex and variegated landscape into a black and white picture.

4.1.3.2 Cannari & Signorini classification

A step forward toward a more sensitive taxonomic approach to clusters localisation has been recently carried out by Cannari and Signorini[458] in a work sponsored by the Bank of Italy. This work starts from the dichotomic Sforzi/Istat classification of SLLS into clusters and non- clusters and introduces a more articulated set of thresholds in order to subdivide the dichotomic classification into 5 classes, which span from super-cluster to non- industrialised areas. The advantages of this approach are multifold. First the classification introduces a multilevel approach to local systems. Second, as Cannari and Signorini point out this may allow the verification of discontinuities in the transition

---

455 (Brusco and Paba 1997) p.279

456 (Brusco and Paba 1997)p.279

457 This is typically the case of industrial clusters in large cities

458 (Cannari and Signorini 2000)

from non- cluster to cluster. Third, the higher sensitivity of the Cannari & Signorini approach to local conditions puts the identification of clusters on a firmer base. However, despite the more sophisticated multicriteria approach, the nature of the identification exercise remains one of finding necessary (but not sufficient) economic conditions for the presence of clusters.

The classification into 5 areas is reported below:

A. Non- industrialised areas (NIA)

B. Industrialised areas of type 2 (A2)

C. Industrialised areas of type 1 (A1)

D. Clusters of type 2 (D2)

E. Clusters of type 1 (D1)

The differences between the categories are based on a multi-criteria scale that include the relative weight of a) manufacturing activities, b) employment in SMEs and c) incidence of specialisation in manufacturing sectors. These are compared against the national average. From A to E the relative weight of the criteria increases. More in details the work by Cannari & Signorini introduces three thresholds corresponding to the three parameters used by the Sforzi/Istat classification. They are:

- Indicator of specialisation α: this is zero if the average of specialisation is less than α times the national average.

- Indicator of employment β: this is used to find the sectors of specialisation. The total of employed population in manufacturing per sector is compared to the national average. If it is β times above the national ratio, the sector is deemed to be specialised.

- Indicator of average dimension of local units (firms) γ: as for β this is used to identify the specialisation sectors. The specific sector is specialised if the average dimension of firms' local units is less than γ times the national average.

Cannari & Signorini have experimented with three values for α (0.5, 1, 2), two values for β (1 and 2) and two for γ (1.5 and 0.7). This gives a total of 12 combinations that have been regrouped in the 5 groups mentioned above.

Non-industrialised areas (NIAs) exhibit the lowest percentage of specialised sectors, the lowest percentage of employment in specialised sectors and the highest average dimension of firms (that is α=0.5, β=1, γ=1.5). At the opposite end of the spectrum (α=2, β=2, γ=0.7) are situated the super-district or D1.

The Cannari & Signorini 's classification respects the conditions indicated in section 4.1.1, reported in the footnote below[459]. First, the grouping into five categories of SLLs describes a set of non-overlapping industrial agglomerations. Second, the range goes from non-industrialised areas to industrial clusters. The former is characterised by the predominance of firms with short range of connectivity (small business serving local markets) and low intensity of connectivity (low phase specialisation and short traditional supply chains). These conditions are more likely to generate an aggregate than a system[460]. At the opposite end sits the highly interconnected and specialised industrial clusters of type 1. Third, the number of sample per category is large enough to allow significant statistical analysis (approximately 400 SLLs belong to the types A1, A2, D1 and D2, the rest – around 350 - to NIA. Fourth, the data are available in a format useful for statistical analysis.

4.1.3.3 A note of caution

Researching on industrial clusters (and in general on self-organising systems) is difficult because by definition the boundary of the phenomenon are not given a priori. Instead they have to be formulated in order to make the research possible. Industrial clusters have no clear boundaries. In fact, they are similar to a swarm of insects orbiting around a nest. The nest is clearly at the centre of the swarm but anyone willing to find the boundary of the colony is clearly at pain to define in a meaningful and unambiguous way the area covered by the colony. It is a matter of degree more than kind.

Given this problem the SLL approach is of interest to this research because it provides an independent definition (and consequent operationalisation) of the boundary of local systems, which is not based on the definition of clusters. This point has several consequences. First, it provides an unbiased approach to the research. As the definition of boundaries is critical to cluster's dynamic, the research can not be suspected of

---

1. distribution of discrete social and industrial agglomerations;
2. distribution of agglomerations ranging from aggregates to self-organising systems;
3. availability of populations large enough to allow statistical analysis, sampling and significance tests.
4. availability of reliable data

[460] see section 2.3.4 for a definition of system and aggregate

adjusting the boundary of the problem in order to make its point. Second, SLLs are (seen from the point of view of cluster) an extremely noisy "container". In fact all types of firms (part of core clusters and not) are recorded in SLLs, thereby 'diluting' the signals coming from the cluster part of the SLL. Moreover, the boundary of the SLL is unlikely to overlap exactly with the boundary of the cluster, ending up either swallowing a cluster in a bigger territory or on the contrary including only a fraction of it. These two aspects of the independently provided definition of boundaries (lack of relatedness to clusters and noisiness of sample) generate an experimental terrain in which the cluster's signal needs to be strong enough to emerge from the background statistical noise. *A fortiori*, they put the eventual results on a firmer ground.

The use of SLL is important from a second point of view. The focus of my research is on the effects of self-organisation in industrial clusters. The method is comparative. Local social/industrial agglomerations (SLLs) are compared against the same set of criteria in order to test whether they cluster according to the probability of including self-organising systems. As the definition of SLLs is independent from (and unrelated with) the definitions of industrial clusters and self-organisation, this method provides an unbiased setting for a comparative quantitative research.

SLL and statistical approach make sense from another point of view as well. Industrial clusters are a very ill defined concept[461]. Several schools[462] exist and consequently there are as many definitions and characterisations as schools. This is almost by necessity, as the cluster organisational form is really a meta-organisational concept, that is, one that exists above the level of formal organisations (often without the members of the cluster being aware of the meta-level). The lack of analytical precision in the studies on clusters depends also on the extreme variability of the cluster form. To counter this aspect, most of the empirical studies have indeed developed the case study approach, thereby contributing to increase the apparent confusion about clusters. The analytical fuzziness of the concept and the multiplicity of cluster forms imply that a direct approach to the statistical study of clusters result very problematical. Being unable to define what the boundary of a cluster is and being equally unable to define the boundary of intermediate

---

[461] (Martin and Sunley 2002)

[462] see section 2.2 for a description of the main schools studying clusters

forms between clusters and non-clusters, the external way to define geographic boundaries, provided by the SLL concept, turns out to be fundamental.

4.1.3.4 The categories of research

The unit of analysis of my research is the SLL in the classification devised by Cannari and Signorini. This ranges from areas with very low industrialisation to areas with industrialisation compatible with Marshallian industrial clusters. The Cannari/Signorini classification of SLLs provides the context for the validation of the propositions reported in chapter 3, which are the object of this thesis. Of the five Cannari & Signorini's categories, the first one (non-industrialised areas) has been excluded from the research for statistical and theoretical problems. The statistical problem has to do with the low numerosity of firms in NIAs. This makes more difficult selecting the NIAs with the minimum required number of firms, as it will be explained in section 4.2.2.2. The theoretical problem is more serious and requires some additional theoretical notions. The whole discussion on cluster relies on the Marshallian concept of economies of agglomeration. Ohlin[463] distinguished between three types of economies of agglomeration:

1. Economies of scale within firms;
2. Localisation economies: these are external to the firm but internal the area and dependent upon the size of the local industry;
3. Urbanisation economies: external to the local industry but internal to the urban area and dependent upon the size of the local economy.

The first type is a well known type of "pecuniary externality"[464]. Localisation economies are invoked to explain geographic clustering of firms in the same industrial sector and dependent supply chain. The causes are the classical Marshallian ones of presence within the cluster of a) pool of skilled labour, b) specialised firms with backward and forward linkages, c) ancillary trades and d) knowledge spillover. Urbanisation economies are similar to localisation economies[465] in having a spatial dependence. However, they depend on the presence of a more complex social, industrial

---

[463] (Ohlin 1933)

[464] This concept expresses the dependence between the price of inputs and size of industry

and institutional system and on economies of scale deriving from the provision of infrastructure, public services, very large and diversified labour market and knowledge spillover. The critical difference between the two depends on who benefits from the economies. In the first case this is the cluster specialisation sector plus the correlated sectors. In the second the benefits are spread across all sectors. The distinction between the two is problematic. In fact, urban areas, which include clusters, exhibit both types of economies. The literature on localisation of innovation presents some evidence that knowledge spillovers are in some cases related to urbanisation economies in other cases to localisation economies. This is especially true for the high tech sector[466].

In order to achieve theoretical clarity I have decided to focus on localisation economies. The focus on localisation economies and the exclusion of urbanisation economies is motivated by the attempt to concentrate on the clusters' dynamic and drivers of economic performance only. Although urbanisation economies are similar in kind to localisation economies, they also hide a complex mix of factors related to the co-location of institutions, firms, infrastructures and services that contribute to the economic performance but are not linked directly to the clusters. Therefore SLLs including large towns have been excluded from the selection of valid SLLs. In general, low industrialised areas with an acceptable number of records (minimum requirement has been set at 600 records, that is firms) are those including large towns. Once towns were excluded from the compute, in the database accessible to the author (see section 4.2.2.1) very few NIAs qualified for analysis.

In short, SLLs are the unit of analysis of this research. SLLs are organised in 5 categories of which the one describing non- industrialised SLL has been excluded from the research.

4.2 THE MEASUREMENT OF POWER LAW IN INDUSTRIAL CLUSTERS

---

[465] In the conclusions I discuss the relationship between localisation and urbanisation economies and diversity

[466] See for the former (Henderson, Kuncoro et al. 1995), for the latter (McCann and Fingleton 1996) and Romer's endogenous growth theory (Romer 1986)

In the previous chapter I have advanced the idea that self-organisation in networks is correlated with a balance between ordered and chaotic features. The distribution of order and chaos tend to follow a power law, which reflects the specific web of flows among nodes across weak and strong links. This is something that a statistical analysis on appropriate systems should be able to test. If self-organisation and power law are correlated, then we can use a power law to test (or reveal) self-organisation in action. From this point of view, a comparison between systems characterised by different degree of self-organisation should reveal a correlation between self-organisation and closeness to power law.

The footprint of a power law is a linear distribution on a double logarithmic scale of the variables characterising the frequency versus the magnitude of a certain phenomenon. Therefore by plotting the chosen variables on a double log graph, a researcher can get an indication whether the phenomenon under inquiry follows a power law or not. Many publications[467] in the literature present the approach by which if something approximating a linear distribution is shown on a double-log scale, then the phenomenon is thought to follow a power law. There is a problem with this approach: the logarithm function. The logarithm allows comparison across different orders of magnitude by levelling out differences in the data sample. Even more so when a double logarithm graph is used to represent data. Therefore the approach of taking the double logarithm approach to represent data exposes the researcher to the risk of seeing a power law where instead the situation could be less clear-cut. Many classes of phenomena, where there is an inverse correlation between frequency and size, will on a double logarithmic scale appear roughly linear. This approach works well if we simply have to discriminate between phenomena described either by a bell shaped curve[468] or by a power law distribution. However, when as in our case we have to discriminate among noisy samples, which are likely to include a mix (maybe continuum or finely graded) of random, hierarchical and scale-free distributions, the above dichotomic approach is not useful.

---

[467] Bak's book (Bak 1997) opened the discussion on self-organised criticality; the popular science book Ubiquity by Buchanan (Buchanan 2000) is a perfect demonstration of the power-law-yes or power-law-not approach. A similar dichotomic approach is found in (Casti 1997).

[468] The distribution of heights of individuals in a population is bell shaped. Most of the individuals will be tightly packed around the average height with rapidly decreasing tails (random networks are characterised by a bell-shaped distribution). In a scale free distribution there is no typical value around which the others values are symmetrically distributed, but the frequency (or numerosity) of events scale inversely with magnitude.

This is exactly the situation concerning my research. Systems of small firms à la Perrow[469] are very diffused in Italy (actually probably more than in any other country in the world[470]). The unit of analysis of this research, the SLL, is a territorial socio-economic unit, which includes all types of organisations and organisational forms: small and large firms, cluster and non- cluster, integrated and dis-integrated companies, service and manufacturing sectors. Because of this dis-homogeneity and of the inherent diversity of systems of small firms, it is quite likely that the statistical differences between classes of SLLs be quite small and unevenly distributed among the classes. In this particular case, the simple semi-qualitative visual approach for the determination of the power law distribution risks to be insufficient. Instead it is more prudent to base one's analysis on a comparison between different classes of comparable phenomena and use a regression analysis to test the closeness to a power law. In this way, no absolute conclusion is reached regarding whether the phenomenon obeys a power law or not, and consequently whether we are dealing with a random or scale-free network, (or whether the phenomenon is critically self-organised or not). On the contrary my approach simply assesses the distance from a perfect power law, in order to establish a correlation between closeness to a power law and self-organising content of the industrial agglomerations. This can be easily done thanks to a regression analysis.

4.2.1 THE CHOICE OF VARIABLES

The first choice regards whether to choose nodes, links or system's dynamic properties to assess a power law distribution. The choice is largely a matter of convenience, if the equivalence between rank-size rule, self-organised criticality and scale-free network approach is accepted. In broad terms, nodal properties are easier to analyse as they are object of economic and census analysis, more than linkages and system's dynamic.

What are the potential candidates for a study of power law distribution in an industrial cluster? I will group them by category:

Node. The obvious candidate is the firm; potential variables are indicators such as firm's size, number of employees, or financial indicators, such as, revenues, profits etc.

---

[469] (Perrow 1992)

[470] (Brusco and Paba 1997)

Links. These are the channels along which exchanges between firms take place. Links could be analysed for instance by mapping the topology of suppliers/customers relationships, using financial input-output indicators.

System's dynamic. For instance the length of customer/supplier transactions induced by an external (to the cluster) order as in the Scheinkman-Woodford model.

Nodal variables are more likely to be available in Census data, whereas the other types may require field research and/or a case study approach. Availability however is only one of the parameters affecting the choice of approach. The statistic data on industrial agglomeration describe punctual data of single firms (profit, employment, birth date, etc), but they do not describe the links among firms. Transactional data exist (for instance input-output tables, see section 4.3.3), but they describe macro-aggregation of transaction between sectors. The information regarding single firms' transactions is lost. The links-based and system's dynamic approaches instead rely on data that have to be acquired via field-research. For instance, one could examine the links in terms of suppliers and customers by applying social network analysis techniques[471]. This has to be done case-by-case and would probably be sensitive to the local aspects of the supply chain, with the implication that comparability of local situations would be hampered by the idiosyncratic aspects of data gathering. Although far from perfect, the nodal approach seems to be the only one that allows effective comparability of data between different types of industrial agglomerations in a statistically significant way. Therefore the nodal approach has been chosen for this part of the research.

In spite of the limitations associated with only analysing nodal properties, the discussion in section 2.4 suggested that the three approaches to power law in networks are equivalent. Therefore the choice of the variable can be done on the basis of convenience alone.

The structural variables selected are number of employees (classes) and number of firms (frequency) pertaining to employee classes. The choice is based on the following reasons:

- availability: data available from Census;

---

[471] (Scott 1991)

- reliability: this type of data is more reliable than financial data[472];
- relevancy: the distribution of number of employees is directly affected (see section 5.2) by the mechanisms of phase specialisation and dis-integration that constitute the idiosyncratic mechanisms at the heart of clusters.

4.2.2 THE PROCEDURE OF CALCULATION

4.2.2.1 Description of data

The data used for this research are taken from the Italian Census of 1991.

The data are affected by some structural problems and limitations:

- There are no data for firms with less than 5 employees. This is a problem for two reasons: first, because the class 1-5 constitutes, in statistical terms, the most numerous class (this is particularly true in Italy – see Table 2); second, because the mechanisms of vertical dis-integration and unbundling, invoked previously, work to increase the weight of micro-firms and small firms in the economy[473]. As the population of firms in industrial clusters is characterised by a very high percentage of micro-firms, the asymmetry introduced in the database by the lack of information regarding the group 1-5 works to make the emergence of the cluster effect more difficult to reveal. This type of error is therefore a conservative error, which does not invalidate the methods used to confirm the proposition.
- The database used for this research results from the merging of two very large databases: the statistical database from the National Institute of Statistics (ISTAT)[474]), which offers aggregated statistical data at a local level and the database from Cerved[475], which instead presents firms specific information. It suffers thereby by inevitable data redundancy and some inconsistencies.

---

[472] There is a general problem of reliability as far as financial data on micro and small firms are concerned

[473] Evans and Wurster (Evans and Wurster 1999) foresee an economy where market mechanisms will regulate the offer and supply of services provided by free-lancers, that is firms of single individuals

[474] See http://www.istat.it/English/index.htm

[475] Cerved is a private company specialising in the provision of professional information for a business use, see (http://www.cerved.com/xportal/home-eng.jsp)

Among the systematic errors present in the database the following one is particularly relevant:

Firms and local units: data in census are based on the distinction between the firm and its units present on the territory. A large company, for instance a corporation such as Fiat, has thousands of local units. Each local unit should be counted in terms of its effective contribution to the SLL, that is, in terms of its financial, employment and other related data characterising the specific area of operation of the local unit. However, these data are unavailable. Therefore each local unit appears in the Italian statistic Census as a full representative of the firm of which it constitutes a part. This means that a Fiat local unit will be reported with (for instance) the full employment figure of the company in its totality. This fact clearly introduces a systematic error which operates in a differential way across the five categories in which SLLs are divided. Let's see why.

For reasons that will be discussed later, the five categories have a different degree of closure to the external environment (see Figure 27). I define closure as the ratio of local units to firms. A closure ratio of 1 indicates that firms have a single local unit. They are usually micro firms. A ratio bigger than one indicates the presence of firms with several local units. This usually indicates the presence of firms of national coverage. Of the four classes, D1 is the most skewed toward the micro-side of the industrial spectrum and toward those sectors where the largest percentage of firms has an endogenous origin. Because the larger the firm, the higher the number of local units, it derives that D1 has the lowest ratio of number of local units to number of firms. This fact is confirmed by the numbers in Figure 27. The other types of SLLS are more affected by this problem. In any case only a relative small percentage of firms show a number of local units bigger than 1 and these firms are largely concentrated in the final part of the tail of the distribution.

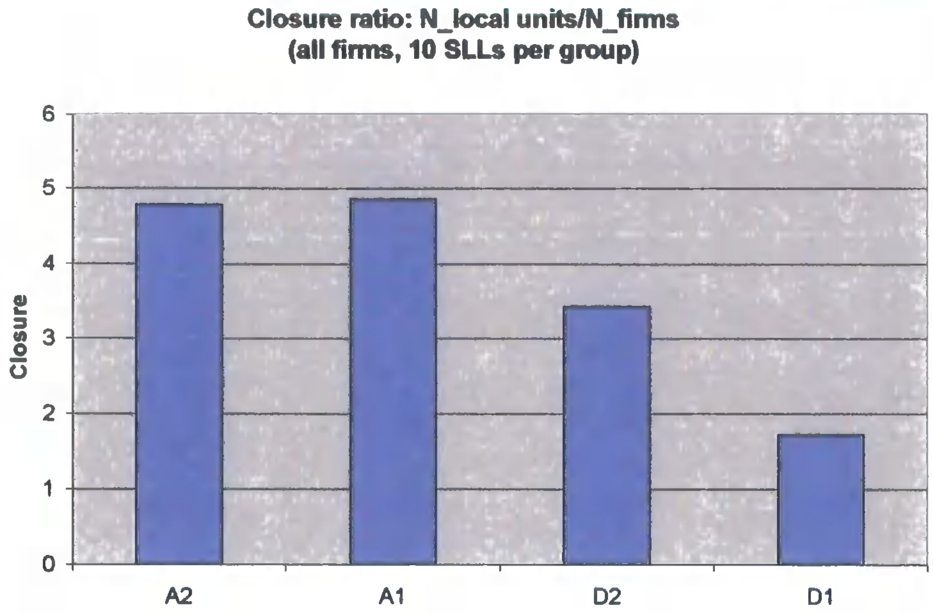


Figure 27: Closure ratio for the categories representing different types of industrial agglomerations

This bias has been corrected by eliminating the final 1% of the firm distribution. This has been defined as the 99% filter. In another attempt to check the influence of this type of error, data have been selected using a filter based on the number of local units as threshold. All firms with $N_{LU} > 50$ ($N_{LU:}$ number of local units) have been filtered out. No significant difference has been observed between the 99% filtered distribution and the N_{LU} filter distribution case. This is due to the fact that the distortion is effective at the very high end of the firm distribution (few firms with hundreds or more of local units). This confirms that the distortion introduced by the incorrect representation of local units in the Census statistics does not significantly affect the analysis of data, once appropriate measures have been implemented to deal with the issue. A further discussion regarding the 99% filter is in section 6.1.2 (cut-off point).

4.2.2.2 Data selection

To get a significant distribution of events for the power law analysis, it is necessary to have a relatively large amount of data. This is very critical in the final part of the distribution (high size, low frequency part of the graph). An initial analysis of power law curve in function of number of events showed a high sensitivity (in terms of the regression parameter) of the curve to a variation in few data in the low frequency part of the graph. To avoid this problem it was decided to adopt a minimum threshold of 600 data per SLL. This means that acceptable SLLs must have a minimum of 600 records (local units). Table 2 shows that the employee class from 1 to 9[476] is disproportionately bigger in Italy than in other main Western economies. The exclusion of the 1-5 class represents therefore a considerable distortion of the population. However, as the error tends to affect SLLs dominated by clusters more than non-clustered SLLs (D1 is more distorted than the others - D2, A2, A1 in order), the validity of the results is not questioned. In fact, since the dynamic of clustering affects micro and small firms more than large ones (for all the reasons traditionally cited in the discussion of Marshallian clusters - see chapters on literature review and theory), the effect of the five employee class truncation will tend to conceal the effect of clustering in the comparative analysis.

Employees Class (% values)	***Italy (1991)***	*Germany (1992)*	*France (1992)*	*UK (1993)*	*Spain (1991)*	*USA (1991)*	*Japan (1991)*
From 1 to employees	**23.3**	**7.4**	**8.1**	**7.2**	**18.3**	**3.0[477]**	**5.0[478]**
10-49	29.2	14.3	17.7	15.6	29.1	NA	NA
50-249	18.9	15.8[479]	21.2	21.7	20.4	NA	NA
1-250	71.4	37.5	47.0	44.5	67.8	36.6[480]	74.1[481]
> 250	28.6	62.5[482]	53.0	55.5	32.2	63.4[483]	25.9[484]
Total	100	100	100	100	100	100	100

Table 2: Class of employees per industry in some OECD countries[485]

---

[476] Census data regarding employees class are usually aggregated around the classes shown above. Consequently only an indicative percentage can be given for the class 1-5

[477] 1 to 10

[478] 1 to 10

[479] 50 to 199

[480] < 500

[481] < 300

[482] > 200

[483] > 500

[484] >300

The database was assembled by the Future Centre (a research centre of Telecom Italia, now part of Telecom Italia Lab[486]) by merging together the Census database from the National Institute of Statistic (ISTAT) with that of a private company (Cerved). Both databases were purchased. Due to confidentiality and privacy issues, I have had access only to a fraction of the database, usually in an input-output fashion. I did have access to specific SLLs but could not browse the whole database. Therefore some selection criteria were adopted in order to guide the data extraction and an appropriate extraction query was constructed.

The selection of the appropriate SLL was based on the following points:

- As explained in section 4.1.2, SLLs are calculated by applying specific statistical algorithms to the Census information. Detailed information on SLLs are published by the Istat[487];
- In 1991 Fabio Sforzi classified dichotomically SLLs as cluster or non- cluster, using specialisation and manufacturing activities indexes. This work was later taken on board by the National Institute of Statistic (Istat) and applied to the 1991 Census data. Since then the Sforzi/Istat approach has become the standard statistical and econometric tool for analysis regarding industrial agglomerations;
- In a little known piece of work[488], two scholars of the Centro Studi (Research Group) of the Banca d'Italia (Cannari and Signorini) have further elaborated and extended the Sforzi/Istat approach, transforming the SLL classification of clusters from dichotomic into a multi-level classification.

The selection was based on the following constraints:

1. Statistical numerosity: to qualify SLLs must have a minimum of 600 entries. The limit of 600 firms per SLL is particularly serious as it reduces the number of useful SLLs. Just to give an example only the upper 20% (around 22 SLLS) of the A1 category fulfils the condition. Excluding from the compute SLLs centred at big cities (3 for the IA1) the useful number becomes less than 20. The situation for the NIAs is even worse: excluding large cities from the compute leaves less than 10 samples.

---

[485] Source: Eurostat: "Enterprise in Europe, Fourth Report, Bruxelles, 1996

[486] http://fc.telecomitalialab.com/english/

[487] (ISTAT 1997)

2. Urbanisation economies: to avoid mixing urbanisation with localisation economies, large cities were excluded from the selection.

3. Populations per category: within the acceptable SLLs, it was decided to randomly select10 SLLs per each category.

4.2.2.3 Calculation procedure

The calculation procedure of the regression coefficients of the size-rank rule is reported below.

1) For each entry (firm) in each SLL, the information regarding the number of employees is analysed. The operations performed on the entries are as follows: first the entries are ranked in terms of size (number of employees). Then frequency of firms per employee classes, cumulative frequency, percentage cumulative frequency, logarithm of frequency and logarithm of bin array (this indicates the employee class, in our case unitary class) are calculated.

2) The auto-filter function is applied to eliminate all record with frequency equal to zero. This is done to avoid conflict with the logarithm function.

3) Percentage cumulative frequency is used to eliminate the final 1% of distribution;

4) Columns with logarithm of frequency and employee class (bin array) are moved to SPSS (SPSS is a very reliable and powerful statistical package whose utilisation has turned out to be necessary due to the repeated failures of Excel to analyse large quantity of data in a reliable way). A regression analysis is performed.

5) All entries (essentially logarithm of bin_array and logarithm of frequency) for every SLL are entered in a large SPSS file, and used as a base for regression and significance test calculations.

Two general types of calculations have been carried out:

A. Regression analysis on a random sample of 600 firms per SLL. This approach is motivated by the attempt to work with samples having the same number of records. This

---

[488] (Cannari and Signorini 2000)

is to avoid any spurious effect due to differences in the number of records. The random selection of the 600 entries from the total is done by using the filtering function in SPSS

B. Regression analysis on the totality of records per SLL. The results for type A and B turned out to be in good accord. In general, method A is much more restrictive and reliable. It usually happens that method B confirms the results obtained by A, amplifying the differences.

4.3 THE MEASUREMENT OF DIVERSITY IN INDUSTRIAL CLUSTERS

Proposition 1 tests the relationship between power law and self-organisation in networks. It is however silent with regard to diversity as a morphogenetic feature of self-organising systems. I have suggested that clusters (as example of self-organising systems) can be explained by means of a set of interdependent autocatalytic loops. Each loop finds its roots in the interaction of diverse parts and stimulates the generation of further diversity. In other words the expansion of diversity in systems is a result of entangled processes. The entanglement is an effect of the connectivity of the system and expresses the idea that the actions of the agents can not be separated from the other agents' actions and from the environment. In one word, systemic diversity and connectivity are inextricably linked. Seen from this angle, Kauffman's rather indistinct definition of the *propagating organisation:* "*...As an average trend, biospheres and the universe create novelty and diversity as fast as they can manage to do so without destroying the accumulated propagating organisation that is the basis and nexus from which further novelty is discovered and incorporated into the propagating organisation*"[489], becomes more precise and subject to empirical analysis. The *propagating organisation* is at any instant of time the product of the processes of exploration times the connectivity that makes those processes possible.

Proposition 2. Self-organising systems maximise systemic diversity (product of diversity times connectivity) at a rate higher than aggregates (and hierarchical systems).

---

[489] (Kauffman 2000) p.85

The rest of this chapter adds some definitional clarity to the concepts of diversity and connectivity and describes the empirical methods used for their calculations.

4.3.1 THE DIMENSIONS OF DIVERSITY

In the discussion on the importance and role of diversity I have treated diversity as an entity without specific attributes. As a matter of fact the empirical characterisation of diversity is problematic. It is intuitively clear that, at least in terms of the science and technology involved, a family of microprocessors, made respectively of three types of semiconductors (silicon, gallium arsenide and germanium - all designed and manufactured according to the physics and engineering of solid state microelectronics) is less diverse than another family of electronic systems whose microprocessors are based on silicon (solid state physics), thermo ionic valves (gas physics) and biological proteins (biophysics). However, if we simply count the number of available options in both cases we get three. On a different dimension, a set of 100 microprocessors, of which 98 are made of silicon and the other two of germanium and gallium arsenide, is clearly less diverse than one in which the options are equally distributed among the three materials. These examples show that the analysis of diversity requires a multidimensional approach.

The problems related with an operational definition of the dimensions of diversity are well reflected in the evolution and history of technology. The basic question is: is the working of evolution witnessed by an increase in diversity in biology and technology?

Most evolutionary biologists support this interpretation. Some like Jay Gould[490] take pain to notice that it depends on the definition of diversity. The so-called Cambrian explosion witnessed the sudden expansion of the number of fundamental phyla (basic body plans). All the known basic body plans were generated in the 'short' span of few million years. Many went successively extinct. In the 550 million years that ensued no new phylum has been created. At the same time the number of genera, families and species present (existing now) is much higher than in the Cambrian. Which period is more diverse: the Cambrian or the current Quaternary? Exactly the same situation is present in many instances of technological history. The beginning of the automotive history saw the competition of three different ways of powering a car: internal combustion engine, steam engine and electrical engine. Few decades later, only the

---

[490] (Gould 1996; Gould 2000)

internal combustion engine had survived. At the same time the number of models, which make use of the internal combustion engine, has skyrocketed. Again, which one is more diverse?

The question boils down to the "*nature or degree of apportionment of a quantity to a set of well defined categories*"[491]. The definition triggers new questions: a) how do we define a set of categories, which can be used as objective criteria to describe the position of the units we are trying to sample/describe on a multi-dimensional diversity space and b) how do we define the lower level units (the indivisible atom or species) of diversity?

Since the Linnean invention of the taxonomic system in zoology, ecology has dealt with the problem of classifying diversity. The Linnean system is based on a series of progressively more inclusive categories (species, families, genera, taxa and kingdoms[492]), which group biological species on the basis of their morphology and reproductive compatibility. Darwinism has introduced an historical dimension to the classification without fundamentally altering it. The fundamental unit is the species. This concept, though not devoid of ambiguities[493], is substantially robust to allow assigning single animals to a species or another, making possible an assessment of total biological diversity.

A robust assessment of diversity is based on three fundamental concepts[494]:

Variety refers to the number of distinguishable actors[495] (otherwise defined as agents) into which a system can be subdivided. It is clear that the capability to determine two individual entities as distinct requires an *a priori* definition of some categories.

Balance refers to the relative density of individual in the different categories.

Disparity instead refers to the difference between the categories themselves. Disparity provides a space where to map species (or the relevant units of analysis) and measure the

---

491 (Stirling 1998) p.38

492 see for instance (Margulis 1998) and (Mayr 2001) p.23

493 there are at least three different ways of defining a species (see (Margulis and Sagan 2002)), which lead to different quantitative assessment of total biological diversity. Other problems with the current definition are for instance in (Mayr 2001).

494 I am using here the excellent work by Stirling (Stirling 1998)

495 (Saviotti 1996)

distance between them. Disparity is important to address the issue mentioned above about the distinction between systems exhibiting the same number of categories.

An important difference between the three dimensions of diversity is the degree with which it is possible to formalise the operationalisation of the dimensions. Variety is a monotonic variable, which grows with the number of "species". If "species" can be defined in a non-parametric way, the calculation of diversity-as-variety does not pose considerable problems. Balance is monotonically dependent on the density of individuals/units in each category of variety. Its calculation can be fairly objective. Disparity instead implies a judgement regarding the relative difference between categories in a certain context. As such, it is an intrinsically qualitative measure.

4.3.1.1 The measurement of diversity

Measuring diversity means developing an algorithm that allows the description (and quantification) of the three aspects of diversity, that is, variety, disparity and balance[496]. Stirling (1998) has developed an algorithm that has the following properties:

Completeness: it includes all three aspects of diversity.

Compactness: it includes a low number of variables.

Transparency: it refers to the minimisation of the number of assumptions behind the algorithm.

Robustness: the algorithm must be as independent as possible from the ordering of the variables used to describe the aspects of diversity.

Stirling introduces an index called M, the integrated multi-criteria diversity index. M satisfies the properties mentioned above. The important characteristic of the algorithm is that disparity is expressed in relative terms as a relative distance between each pair of entities in an *ad hoc* introduced space, called the disparity space. The idea behind the disparity space is simple. The calculation of diversity as I have mentioned above is characterised by the inevitable judgemental and context-dependent assessment of disparity. As disparity is an indication of the distance between the categories among which the units are apportioned, the calculation of disparity necessitates a set of criteria to organise the categories in a disparity space. Once this is unambiguously given, the

---

[496] (Stirling 1998)

distance can simply be calculated as the geometric distance in a Euclidean space between the positions of each pair of units. Diversity becomes then the sum over all possible pairs of the distance between each pair of units multiplied by their relative frequency.

Equation 7[497] $$\mathbf{D = \Sigma_{ij}\, d_{ij}\, p_i\, p_j}$$

The suffixes *i* and *j* indicate two generic units (or species), d_{ij} represents the distance between *i* and *j*, $p_i\, p_j$ stands for the product of the balance, that is number of units of species *i* and species *j*. This simple equation captures all aspects of diversity.

4.3.1.2 The Measurement of connectivity

Assessing the diversity of a portfolio of options, or the elements of a set, is useful for a static assessment of the degree of dispersion of some properties along the axes of diversity chosen as representative. But it is less relevant for the point of view adopted in this thesis: that is, that the study of diversity in interconnected systems shows generative properties. A measure of diversity that does not address the interactions between parts is not able to distinguish between a system and an aggregate. This negates the possibility of using the measurement of diversity as a tool to assess whether diversity itself has autocatalytic properties. Assessing interconnected features requires an algorithm that links diversity and connectivity. A simple way, again suggested by Stirling[498], consists in introducing in Equation 7 a term indicating the relationship between the species for each pair of species under examination. The form of Equation 7 makes this straightforward. Indeed it is enough to introduce a multiplicative term c_{ij}, which captures the presence and degree of interaction between each pair of elements. The effect of the new index is that of excluding isolated elements from the compute of diversity. In fact, if the elements *i* and *j* are isolated from one another, the resulting connectivity measure c_{ij} will be void. As the new index measures diversity in interconnected sets, I will define the new index systemic diversity and indicate it with D_s

Equation 8[499] $$\mathbf{D_s = \Sigma_{ij}\, d_{ij}\, p_i\, p_j \cdot c_{ij}}$$

---

497 (Stirling 1998)

498 (Stirling 1998), p.110

499 (Stirling 1998)

4.3.2 THE MEASUREMENT OF GENERIC DIVERSITY IN INDUSTRIAL CLUSTERS

What I develop in the following part is an original way to operationalise Stirling's algorithms in order to measure diversity and connectivity in the context of industrial agglomerations. The measurement of diversity in industrial clusters requires an operational definition of variety, disparity and balance. I define variety in terms of firm's sector by using SIC codes. Disparity is defined in terms of a unidimensional distance between SIC codes, in a way I am going to explain better below. Balance is simply given by the number of firms populating each specific niche (or SIC code).

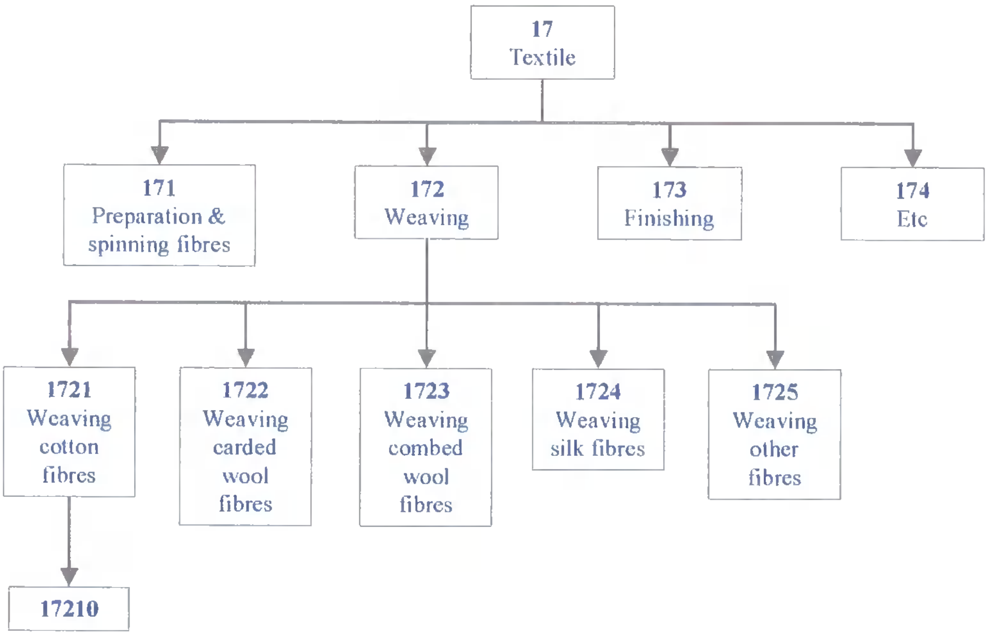


Figure 28: Structure of Ateco codes (detail of textile branch)

SIC codes are known in Italy as Ateco codes. The Ateco are organised as in Figure 28.

The Ateco codes introduce a set of criteria to classify firms on the basis of their industrial specialisation. Specialisation is defined according to the typology of products, technology and/or process used. As the classification was originally devised for an economy dominated by tangible goods, it works better for the manufacturing sector than for the service sector. The interesting point for the purpose of this thesis is that SIC codes identify specific industrial 'species' by positing them into larger groups. These groups, like in a Linnean taxonomy, share some taxonomic traits. An example of this tree-structure is shown in Figure 28. The first digit distinguishes between manufacturing and service sector. The second introduces sub sectors within the ones previously identified. The code 17, for instance, refers to the textile industry. Within the code 17, 171 identifies preparation and spinning of fibres, 172 weaving of fibres, etc. The 172 constitutes in its turn a family including the sub-sub-sectors of weaving cotton (1721), carded wool (1722), combed wool (1723), etc. As in a Linnean tree, the classification goes from more inclusive groups (fewer digits) to less inclusive ones (more digits). The maximum resolution, the one achieved with the maximum number of digits, identifies the unit of diversity measurement, that is, an industrial species.

Once the context of the diversity measurement is stated in these terms, the total variety of a specific cluster is simply given by the number of species, that is by the finest resolved number of SIC codes. The assessment of balance instead requires a calculation of the density of firms per unit of variety, that is, SIC code.

The assessment of disparity is more complex. This requires a calculation of distance between the categories of variety, that is, the SIC codes, taking into account their level of aggregation in the tree structure. A way to calculate the distance is to use the biological taxonomic metaphor. In biology two species are described as belonging to the same family, say, Canis familaris (common dog) and Canis Lupus (wolf), when they share a common progenitor, the originator of the family. In turn, two families will belong to the same genus (the next level of aggregation), when they have again a common ancestor, the originator of the genus. Clearly, the case of two species belonging to the same family will exhibit a smaller disparity than the case of two species belonging to different families. By the same token the distance between species belonging to families not related at the genus level will still be larger. In order to measure disparity I propose to use the number of bi-furcations along the SIC code tree structure: the higher the number of bifurcations between species, the larger the distance. In other words, because all species belong to progressively more inclusive groups, the distance between any two species simply reflects the number of steps separating the species from the common least inclusive group (minimum common denominator).

4.3.2.1 Limitations of taxonomic approach

The taxonomic approach to the calculation of diversity of industrial species presents some risks and limitations. These are discussed in this section.

First, there is an issue regarding divergence and convergence in taxonomic trees. The taxonomic approach works satisfactorily well in biology because the speciation process is assumed to be always divergent[500] and cumulative. In technological evolution instead convergence between different technological lineages or species takes place. This makes the research of a common progenitor less straightforward.

Second, the definition of species is ambiguous. How are we for instance to assess the distance (in absolute terms) between two species using different processes (say weaving and finishing) on the same fibre (e.g. cotton) from two species using the same process over two different fibres? This objection is valid in the case of an absolute measure of diversity being required. The Stirling algorithm, however, only measures relative distances averaged over all possible pairs in a cluster. Therefore the identification of a species becomes less problematic, as relative distances will be less sensitive to the definition of species, provided that the distance between pairs are calculated according to the same criterion.

Third, the measure of distance reflects industrial, technical or product difference not supply chain distance. SIC codes give an indication of specialisation based on a set of criteria that reflect type of product and technique of production. As such a high disparity (large distance) does not automatically imply anything in terms of potential supply chain closeness. For instance, a firm manufacturing electronic components is separated by many steps from a firm specialising in weaving cotton. However, the former could be a direct supplier of the latter. This point is not critical for the calculation of systemic diversity. In fact, first this approach tends to under estimate the distance in SLLs dominated by highly connected clusters (in which a higher percentage of firms will belong to sectors closer to the cluster's dominant specialisation) than in SLLs where the pattern of industrial activities is more evenly spread across all possible SIC codes. As my thesis claims that clusters are more systematically diverse[501] than non-clusters, the

---

[500] at least the NeoDarwinist assume so. Recent evidence shows that at least in the bacteria world and most probably in the fungi, plants and animals kingdoms as well speciation by convergence of different lineages through the process of symbiogenesis is possible (see (Margulis and Sagan 2002)).

[501] Systematic diversity refers to product of diversity times connectivity

method of calculation of disparity, if anything, requires a stronger signal to emerge from noise. Second, the calculation of connectivity will provide an independent measurement of interaction among species, thereby reinstating a more realistic assessment of distance.

Fourth, identification of firms by means of SIC code is problematic. I assume for the sake of computability that firms are uniquely described by a single SIC code. This is largely true for small firms, but clearly less so for larger firms. This error tends to over-estimate diversity in SLLs dominated by a larger percentage of small firms.

Fifth, appropriateness of SIC codes to study clusters. SIC codes are an appropriate tool to describe slowly changing situations, whereby cumulativity of effects creates clearly demarcated industrial niches. Instead, where market turbulence and technology effervescence rapidly modify industrial structures, either by introducing new industrial species or changing the features of the existing ones, it is difficult for SIC codes to keep up with the changes. This is especially true for industrial clusters which represent a meta-stable organisational form, whose survival requires a constant renewal of organisational process, products and technologies. Moreover, the processes of phase specialisation and vertical disintegration generate a very granular industrial texture. As the clusters' organisational structure presents idiosyncratic aspects, these are not properly accounted for by methods of classification based on national averages, such as the SIC codes. In other terms, the description of the clusters' industrial structure by means of SIC codes will average out the clusters' differences. This is also a considerable source of error, but because it tends to under-estimates the proposition (it fails to account for additional diversity), it will not detract from the results.

4.3.2.2 Procedure for calculating generic diversity

The calculation of diversity is carried out as follows. Although conceptually simple, the implementation of the calculation requires some attention and some dedicated programming. In particular SPSS has been used for statistical analysis and MatLab for dealing with operations concerning large matrices.

The first step, that is the calculation of generic diversity requires the following steps:

A. calculation of d_{ij}, distance between SIC codes;

B. calculation of ***D***, generic diversity for SLL *k*. This is done in two steps: first the product ($d_{ij}\, p_i\, p_j$) of distance times frequency of occurrences (number of firm in SIC niche) is calculated. The product yields the generic diversity for a single pair of SIC codes. Second, the sum over all possible pairs of SIC codes leads to the value ***D*** ($D = \Sigma_{ij}\, d_{ij}\, p_i p_j$), which represent the generic diversity for SLL *k*.

C. a statistical analysis of generic diversity across SLL categories is carried out.

A fuller description of these calculations is given below.

A. The assessment of distance hinges on finding the point of separation at the minimum common denominator between two SIC codes in the SIC code tree structure. A simple procedure for the implementation of the distance calculation is as follows. The 5 digits SIC codes describe from left to right progressively more detailed and less inclusive groupings. This fact indicates that the weight to attach to distance decreases from left to right. In fact, two firms that differ at the first digit are more diverse than two firms that differ at the final one. This aspect can be captured in terms of bifurcations, which represent the points where the tree structure branches out. Two firms that differ only for the last digit are part of the same tree structure until the final bifurcation. Firms that differ at the first digit instead follow different paths since the first bifurcation. Therefore it makes sense to stop the analysis of distance at the first different digit (the most important one). As there are five possible bifurcations, the distance between any two SIC codes comes to be one of five possible types. The distance between any two SIC codes can then expressed by means of a string of 0s and 1s, whereby the position of 1 indicates the level at which the firms diverge. An important decision concerns the weight to attach to each difference. It is proposed that the weight goes with the power of two: difference at 5^{th} digit is 2^0, difference at 2^{nd} digit is weighted with 2^1, etc. To test the sensitivity of the calculation on the choice of weight, the calculation has been repeated using linear weight (from 1 to 5). No significant difference has been observed.

a. The detailed procedure involves the separation of each SIC code into a string of digits. This is done in order to analyse each digit in isolation. This is achieved by the first part (called digitisation) of a MatLab program, written for the occasion, shown in Appendix .

b. For each entry (firm), the bifurcation point along the SIC tree structure is found and the relative weight is attached to it. The value is inserted into an n x n matrix, called the distance matrix or ***dis*** (in the program this matrix is called ***K)*** . Each element of the matrix contains the distance between the firm ***i*** and the firm ***j*** of the cluster.

c. The sum over all pairs is carried out and a numerical value expressing the total distance between all activities (recorded in the database) of the cluster is obtained. This number gives an idea of the disparity of the species in the SLL. The programming codes for points b and c is shown in the MatLab program under the heading Difference.

B. The calculation of generic diversity is more straightforward. For reasons of compatibility with the connectivity calculation (see next section), only the first three digits of the SIC codes have been assumed as significant. This step requires the matricial calculation of dis_{ij}, distance matrix, times the 1 x n matrix (vector) p_i expressing the density of firm population across the industrial SIC niches. The vector matrix is simply calculated using SPSS. Each $dis_{ij} \cdot p_i \cdot p_j$ is stored as an element div_{ij} of the generic diversity matrix ***div***. Summing over all cells of the matrix ***div*** gives the generic diversity characterising cluster *k*. Results and lines of program are shown in Appendix (the specific code for part B is shown under the heading *calculation of diversity index*).

C. This part analyses average, standard deviation and significance of results of generic diversity over all clusters subject of analysis.

4.3.3 THE CALCULATION OF CONNECTIVITY

A cluster is an area where deepening of specialisation is accompanied by a multiplication of linkages among cluster's agents. The density and nature of linkages (strong and weak links) is higher and different than the traditional forward and backward linkages of traditional supply chains. In fact one of the reasons that the literature advances to explain clusters' higher productivity, competitiveness and resilience is higher interconnectedness[502].

Linkages (between firms) can take a variety of forms. For the sake of this work, I refer to stable transactional linkages in supply chains. The problem that arises is how to reveal those links and assess their intensity and nature.

---

[502] see for instance (Porter 1991)

The first way is to directly inquire into them by adopting, for instance, a snowballing technique[503]. This consists in reconstructing the web of links by connecting a node to all its neighbours, then by connecting the neighbours to their neighbour's neighbours, etc. The number of links diverges quickly, limited only by the dimension of the sample and the degree of resolution adopted. This technique is effective and sensitive to local circumstances, although it depends by default on the respondents' understanding of what is meant by a linkage. It is also extremely time consuming, especially if the researcher is looking for a high level of resolution. It is a bottom-up approach with the typical advantages and disadvantages of inductive methods of field research.

The alternative approach consists in using financial data and in particular economic input-output tables. These provide a quantitative indication of the transactional intensity within and across industrial sectors. The immediate advantages are comparability and easiness of implementation. The disadvantages instead concern the loss of adherence to (and therefore understanding of) local circumstances and consequently the calculation of generic values of connectivity (obtained manipulating data from aggregated regional or national averages).

The latter approach has been adopted for ease of comparability and availability of financial data. The use of input-output (I-O) tables for analysing industry linkages is well tested in literature and sophisticated algorithms are available[504]. These analyses focus on the reconstruction of buyer-supplier links in order to assess the degree of embeddedness of regional economies[505]. I-O tables give an indication of the transaction intensity between two industrial sectors by calculating the financial interactions that occur between them.

I have made use of the I-O tables[506] published by the National Institute of Statistics (ISTAT) in 1992. These tables provide a quantitative indication of the financial interconnection between industrial sectors[507]. The use of I-O tables in the analysis of connectivity has some disadvantages. First, the classification of industrial sectors is rather

---

503 see for an introductory treatment (Scott 1991)

504 (Chenery and Watanabe 1958; Rasmussen 1958; Dietzenbacher 1992)

505 (Midmore, Munday et al. 2002)

506 (ISTAT 2000)

507 The sectors are classified according to the European industrial code classification NACE70, which is slightly different from the SIC – Ateco in Italy – used in the course of this research.

course, with manufacturing and service industries divided in 118 categories, making the analysis of intra-industry linkages less defined than the analysis of inter-industry linkages. Second, the national dimension is imposed on all local agglomerations, as if the rules of interactions between sectors were geographically independent. Third, the *one size fits all* approach neglects the fact that industrial clusters (because of the disintegration induced by phase specialisation, network reconfiguration and high rate of product and organisational innovation) do not properly follow generic industrial classification, which are instead more appropriate for other areas of the country and for more traditional industrial settings. Fourthly, the different classification between SIC and I-O tables imposes a conversion of one into the other. However, imposing the straightjackets of national averages is more likely to under-estimate the phenomenon sought than vice-versa. In fact, if the original forms of transactional relationships between firms characteristic of industrial clusters are not captured by the SIC codes and I-O tables, then the overall effect will be the under-representation of the cluster's diversity and connectivity. As under-estimation is a type of error skewed toward conservatism, this does not compromise the validity of the results.

4.3.3.1 Procedure for calculation of connectivity

This involves the calculation of a connectivity index given by the sum of an index of financial transactions over all pairs of firms in an SLL. The steps to calculate the index are the following:

A. Harmonisation of I-O codes with SIC codes. This yields a general I-O matrix C. The I-O code classification of industrial activities does not follow the more general SIC classification known as Ateco, but instead the European NAVE70 (in general the NAVE70 is very close to the ATECO, as it represents a spin-off of the ATECO, introduced to harmonise the Italian system with the European practice). This classification falls in between two and three Ateco digits; it is less comprehensive than the 2 digits but more comprehensive than the three digits. The harmonisation between the two classification procedures has been carried out by expanding the I-O matrix into the three Ateco digits. This has been done by proportionally splitting the more inclusive 2.5 I-O codes into the more resolved three Ateco digits.

B. Extraction of a sub-matrix C_k. This is carried out by extracting from the Ateco I-O matrix C the rows and columns that correspond to the industrial species present in cluster k. This provides a sub-matrix called C_k. The matrix C_k contains the SIC codes relative to cluster k. The Mathlab programming code is shown in Appendix 1.

C. Sum over all elements of C_k gives a numerical value proportional to the connectivity of cluster k. The sum over all elements ***i and j*** of matrix C_k gives a quantitative indication of the connectedness of a cluster.

4.3.4 THE MEASUREMENT OF SYSTEMIC DIVERSITY

The algorithm for the calculation of systemic diversity is given **Equation 8** below.

$$D_s = \Sigma_{ij} d_{ij} p_i p_j \cdot c_{ij}$$

The procedure for the calculation of D_s is the following:

A. For each cluster k the elements i, j of the matrix C_k are multiplied by the same element of the matrix of generic diversity div_{ij}. This constitutes the element i, j of the matrix D_s.

B. The sum over all i, j of the product $div_{ij} \cdot C_{ij}$ gives a numerical value which is an indicator of SLL k's systemic diversity. The lines of programming are shown in Appendix . This part of the program is called: calculation of systemic diversity index.

The statistical analysis of distance, generic diversity, connectivity, and systemic diversity provides an indication regarding the falsity or truth of the proposition concerning diversity and connectivity

Chapter 5: Results

5.1 STRUCTURE OF CHAPTER

This chapter introduces the main empirical results relative to the demonstration of the two propositions derived in the previous chapter. The first part illustrates the evidence concerning the relationship between self-organisation and power law in the context of industrial agglomeration (the first proposition). The second part introduces the evidence regarding diversity, connectivity and self-organisation (the second proposition). The results are discussed in chapter 6.

5.2 POWER LAW RESULTS

The proposition regarding power law and self-organisation is reported below:

Proposition 1. Self-organising systems are closer to a power law than aggregates (collections of parts) or hierarchical systems. For industrial agglomerations (SLLs), the closeness to a power law is related to the degree of self-organisation present in the agglomeration, and emerges in the agglomeration's structural and/or behavioural properties subject to self-organising dynamic[508].

The proposition assumes that the nature, intensity and configuration of relationships between organisations in industrial clusters affect some properties of the agglomeration, thereby making them different from other types of industrial agglomerations. As the nature of the difference is dispersed across a population of firms and is due to the systemic effect of the self-referential properties of the web of linkages, it implies that the appropriate level of analysis is the population and not the single firm. Therefore, the approach followed is a statistical one applied at the level of population of firms. There is a further aspect related to the web of linkages. The set, nature and range of linkages determine whether the system behaves as an aggregate (i.e. a random network) or system (i.e. a non-random network). The footprints of the two types are different. The former is characterised by a bell distribution of the relevant properties, the latter, in the case of self-organising networks, by a power law. My analysis in section 3.2 suggests that a

[508] The various concepts mentioned in the proposition are discussed in section 4.1.2

power law analysis can be carried out at the nodal, link, and dynamic properties of networks. In fact the three sets of properties are linked in such a way that if a power law behaviour emerges in one set, then it can be assumed to be valid for the other two.

The presentation of the results is divided in two parts. Section 5.2.1 gives more details concerning the approach I have implemented to calculate power laws in industrial agglomerations and validate the choice of the regression analysis. Section 5.2.2 presents the findings regarding power law distributions.

5.2.1 REGRESSION ANALYSIS

The objective of this section is to develop a robust method to test the power law proposition. This is done in the following way. Ten samples from each category (A1, A2, D1 and D2) are grouped in a single SPSS file and the following calculations are performed via SPSS on each category. Regression analysis has been chosen as it permits to determine the probable form of relationship between two sets of variables. In order to apply regression analysis, a series of assumptions[509] have to be satisfied:

1. The independent variable X to be measured without error. This is verified as the chosen variable X represents the spread of number of employees. The same set is used in all sub-populations of Ys.
2. The sub-populations represented by the dependent variable Y are normally distributed for each set of values of X. This has been demonstrated by means of residual analysis[510]. In fact, as it is shown from Figure 29 to Figure 32, the distributions of residuals in the four classes of data are approximately normal.
3. The variances of the sub-populations Ys are approximately equal. Results are shown in Table 3.
4. The means of the sub-populations Y are approximately equal. See Table 3.

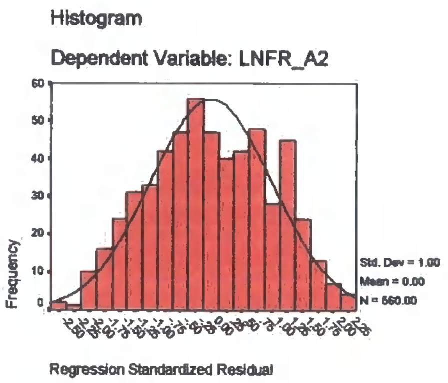
[509] (Daniel and Terrel 1989)

[510] (Daniel and Terrel 1989) p.481

	A2	A1	D2	D1
Mean	3.78	3.67	3.60	3.50
Variance (sum of square: regression + residual)	931	1085	1096	1080

Table 3: Means and variances of sub-populations

The second logical step concerns a visual check of the linearity of the scatter of data (logarithm of size and frequency)[511]. This is shown from Figure 33 to Figure 40. Even a visual check of the scatter plots for the four categories confirms that the distributions seem to be linear and therefore strongly supports the application of regression analysis.



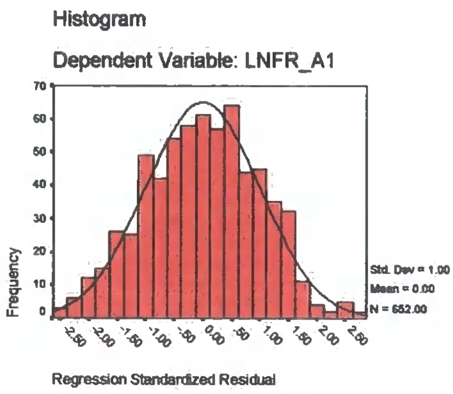
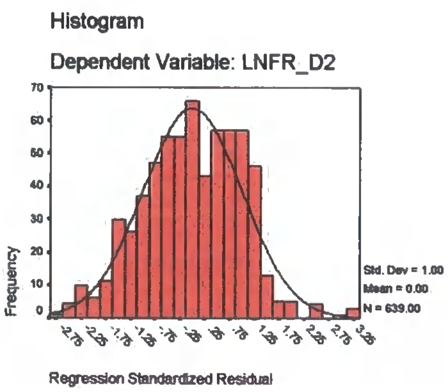
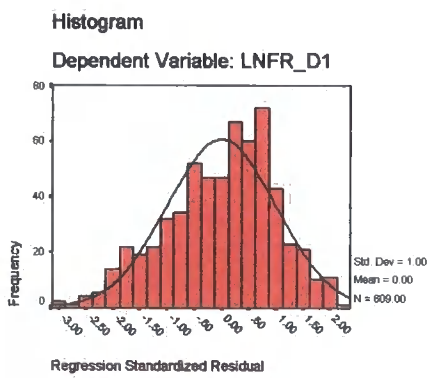


Figure 29: Distribution of residual errors for A2

Figure 30: Distribution of residual errors for A1





[511] The symbols in the graphs are the following: the class indicates the SLL type, Ln_Array represents the logarithm of the employee dimensional class, and ln_frq the number of firms in the corresponding dimensional class.

Figure 31: Distribution of residual errors for D2

Figure 32: Distribution of residual errors for D1

Are the four categories representative of four different populations or are they drawn from the same one? A way to answer this question involves the comparison of populations' means and variances. This can be done by making use of the F test. The null hypothesis is that means and variances across categories are the same within random fluctuations limits. Table 4 shows the analysis of variance (Anova).

ANOVA

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	22.010	3	7.337	4.297	.005
	Residual	4193.524	2456	1.707		
	Total	4215.534	2459			

a Predictors: (Constant), D2, A2, D1

b Dependent Variable: LNFR

Table 4: Anova analysis

We can use the table of F-test to find the threshold below which H_0 is rejected. For a significance level of $\alpha = 0.005$, the threshold is 4.28, which is smaller than F = 4.30 reported in Table 4. Therefore the null hypothesis is rejected. In fact, the chance that the null hypothesis is true is less than five per thousand. The samples do not come from the same population.

Two points have still to be clarified. First, are all categories different from each other (samples drawn from different populations)[512]? Second, how to estimate the degree of closeness to a power law of the four categories?

The first question can be answered by making use of the univariate analysis of variance. In this test each category is compared with the other three and an assessment of their population difference significance is given. Table 5 shows the results for D1 (comparison between D1 and A1, A2 and D2). The important column is the significance

[512] the previous analysis simply demonstrates that not all the samples are taken from the same population, not that all of them are respectively different from all the others

test (indicated as Sig in the table). This shows that the null hypothesis is denied for D1-A1 and D1-A2 for a significance level smaller than 5%. The null hypothesis is not rejected in the case of D1-D2, at least not with 95% accuracy. As D1 is the outlier of the four categories, the other three tables, obtained by comparing A1, A2 and D2 with all the other categories, do not reveal any new aspect (see Table 6 and Table 7)

Parameter Estimates

Dependent Variable: LNFR_A1

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	1.262	.053	23.832	.000	1.158	1.366
[CLASS=A1]	-.232	.074	-3.150	.002	-.376	-8.757E-02
[CLASS=A2]	-.226	.077	-2.954	.003	-.376	-7.598E-02
[CLASS=B2]	-.114	.074	-1.540	.124	-.259	3.113E-02
[CLASS=D1]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Table 5: Significance test for D1 (D2 renamed as B2 for software reasons)

Parameter Estimates

Dependent Variable: LNFR_A1

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	1.148	.052	22.207	.000	1.047	1.249
[CLASS=A1]	-.118	.073	-1.622	.105	-.261	2.464E-02
[CLASS=A2]	-.112	.076	-1.481	.139	-.260	3.630E-02
[CLASS=D1]	.114	.074	1.540	.124	-3.113E-02	.259
[CLASS=D2]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Table 6: Significance test for D2

Parameter Estimates

Dependent Variable: LNFR_A1

Parameter	B	Std. Error	t	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Intercept	1.036	.055	18.760	.000	.928	1.144
[CLASS=A1]	-5.97E-03	.075	-.079	.937	-.154	.142
[CLASS=D1]	.226	.077	2.954	.003	7.598E-02	.376
[CLASS=D2]	.112	.076	1.481	.139	-3.630E-02	.260
[CLASS=F2]	0[a]	.	.	.	.	.

a. This parameter is set to zero because it is redundant.

Table 7: Significance test for A2 (renamed as F2)

The final point has to do with the assessment of closeness to a power law. This analysis is carried out by using the test known as r^2. The coefficient r^2 gives "*a measure of the linearity of the data point*"[513] or "*a measure of the closeness of fit of the regression equation to the sample data*"[514]. This implies that we can use r^2 as an indicator of closeness to a power law of the distribution of records in the four categories. As we have already demonstrated that the populations are statistically different, especially in the case of D1, this allows us to infer that the eventual differences between the four categories in terms of closeness to a power law are likely to be robust.

5.2.2 THE POWER LAW RESULTS

The demonstration of the validity of the proposition is shown below. The first thing to do is to analyse whether populations of firms in single SLLs follow or not a power law. A sample of four SLLs relative to the four classes under analysis is shown from Figure 33 to Figure 36. As discussed in section 4.2.1, the chosen variable for power law analysis is the firm size (number of employees). The figures below report the logarithm of firm size expressed in terms of number of employees (LNARR – logarithm of the array) against the logarithm of the number of firms with that size. The figures show that small firms (with a low number of logarithm of the array) are much more numerous than firms with a high number of employees. The figures also show that larger firms are often represented by a single entry, thereby yielding the long horizontal line corresponding to the logarithm of frequency equal to zero.

---

[513] (Daniel and Terrel 1989) p.481

[514] (Daniel and Terrel 1989) p.480

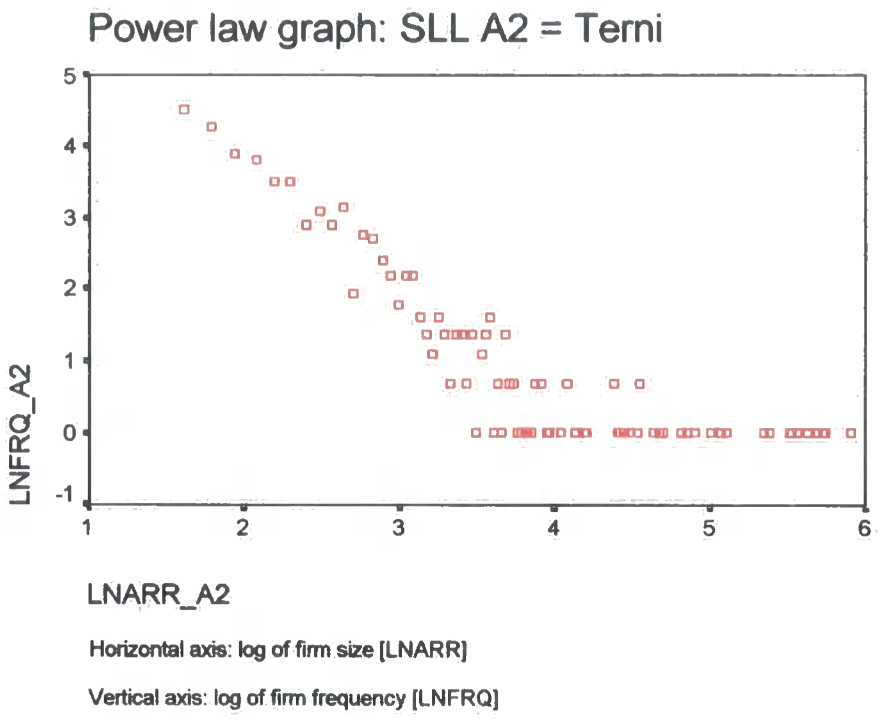


Figure 33: Scatter graph of power law for A2 type SLL. The horizontal axis represents the logarithm of firm size (number of employees or LnArray), the vertical axis shows the logarithm of the corresponding frequency (lnFrequency)

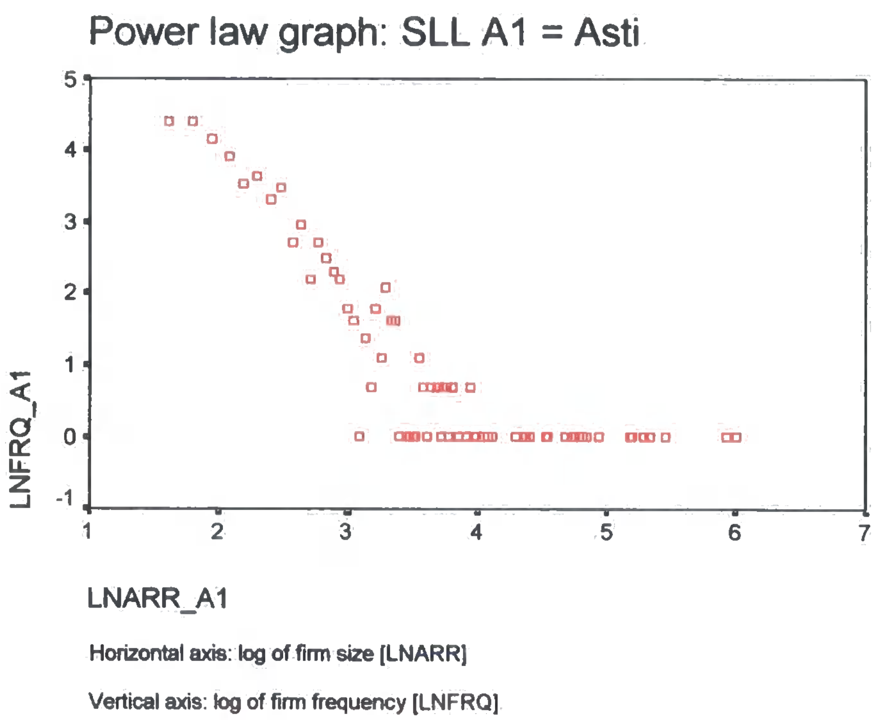


Figure 34: Scatter graph of power law for A1 type SLL

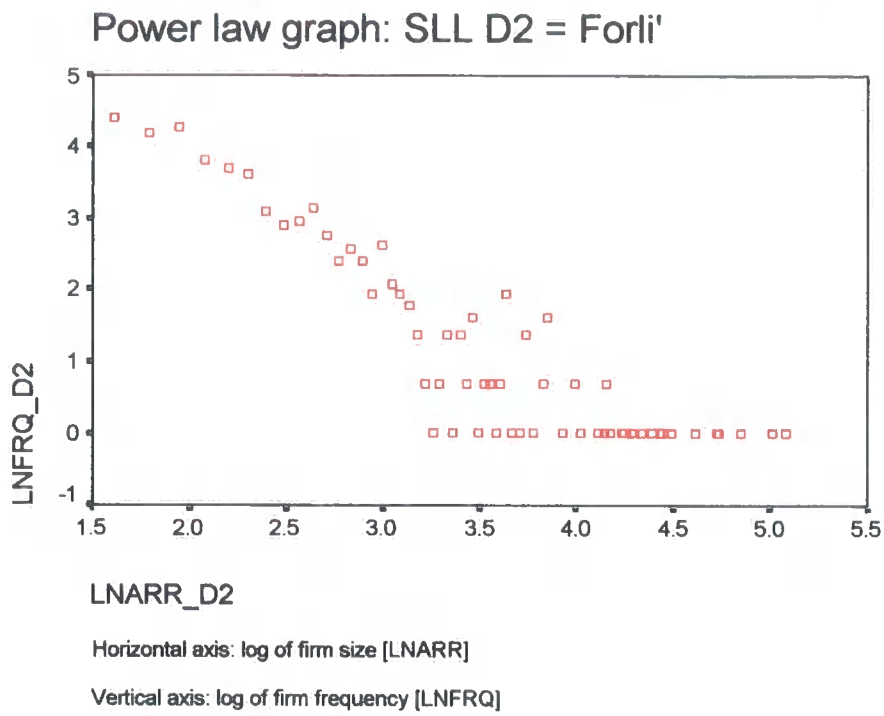


Figure 35: Scatter graph of power law for D2 type SLL

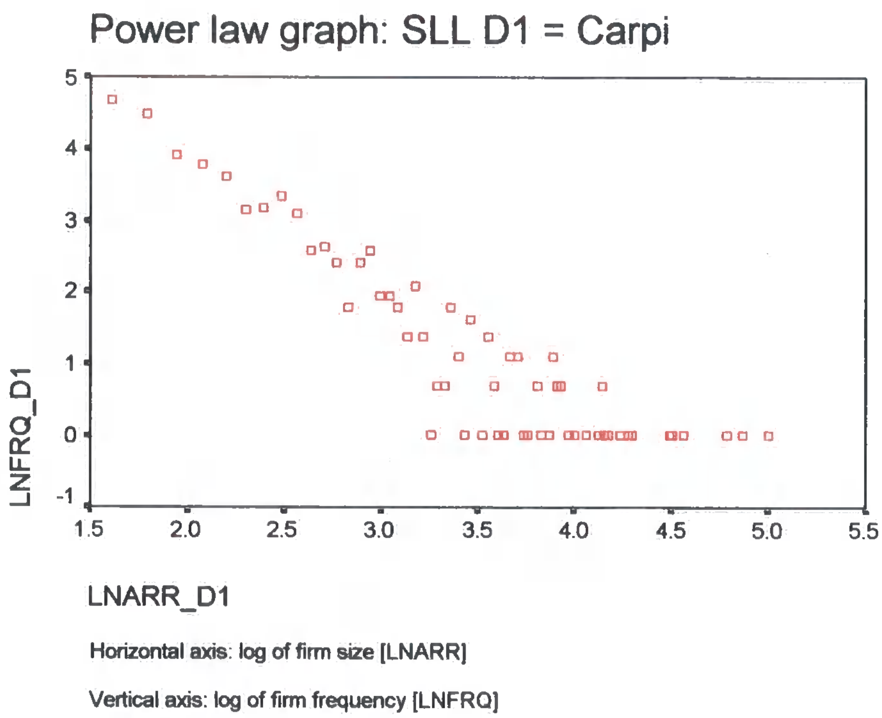


Figure 36: Scatter graph of power law for D1 type SLL

The figures show a typical power law distribution: very broad range of dimensional sizes, absence of a typical scale of the distribution, high frequency of small events (small firms) and low frequency of large events (big firms), linear relation between the distribution of sizes and frequency on a double logarithm graph. The figures show that irrespective of the SLLs class, the closeness to a power law is very marked in all four examples. This is somehow surprising, as it seems to deny the relationship between self-organising properties of clusters and power law.

A single SLL per class is not statistically representative. The next level of analysis tests the proposition at the level of a population of populations. In section 4.1.3.2 I introduced the Cannari & Signorini classification, in which SLLs are grouped according to the probability of including a Marshallian cluster. In my analysis I have analysed ten SLLs per category on which a power law analysis is performed on the samples. The results are shown in the figures below.

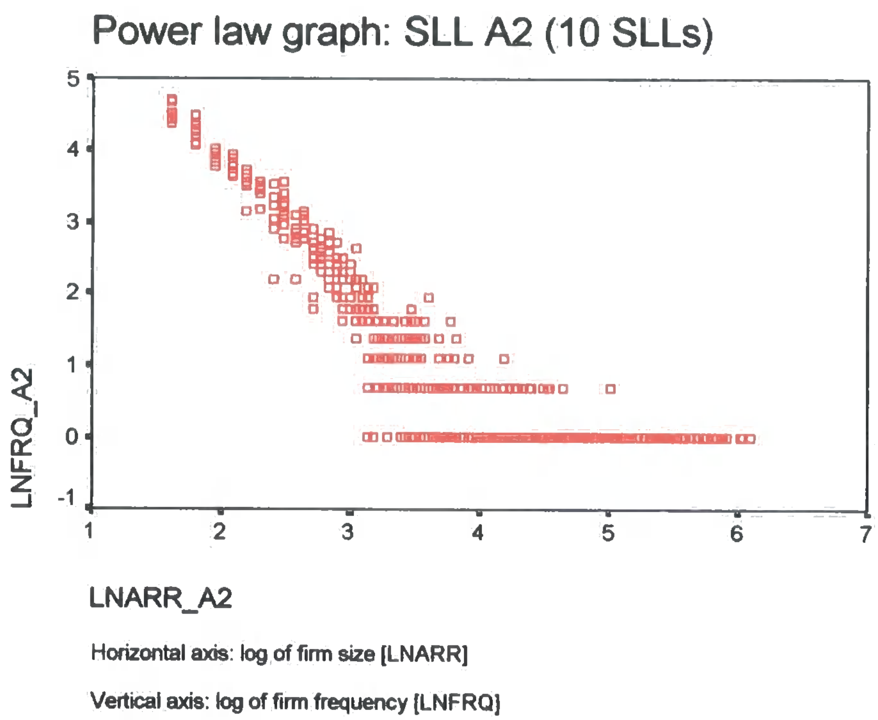


Figure 37: Scatter graph of power law for A2 (10 SLLs), The horizontal axis represents the logarithm of firm size (number of employees or LnArray), the vertical axis shows the logarithm of the corresponding frequency (lnFrequency)

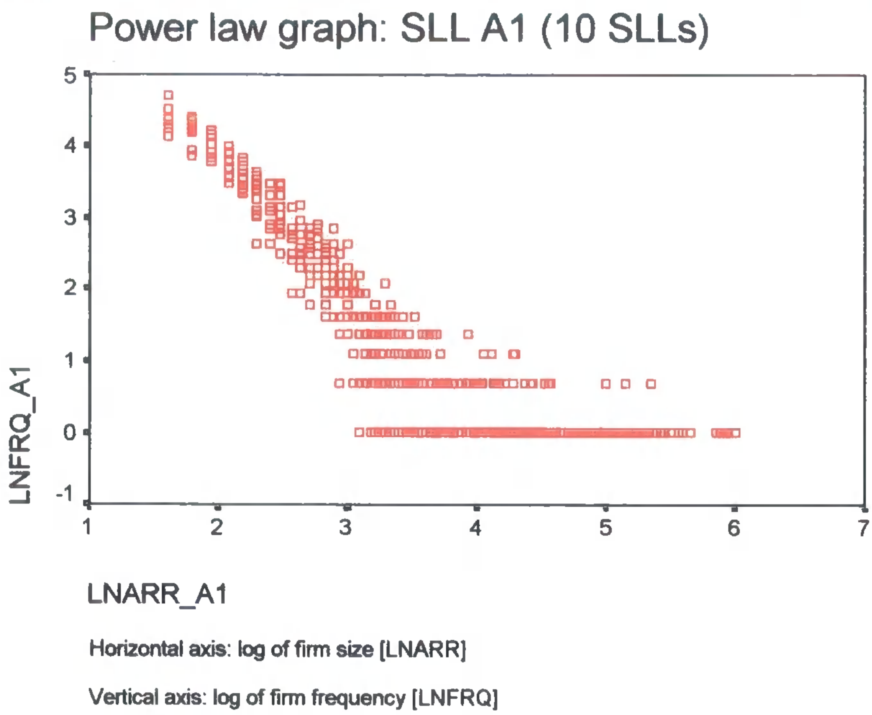


Figure 38: Scatter graph of power law for A1 (10 SLLs)

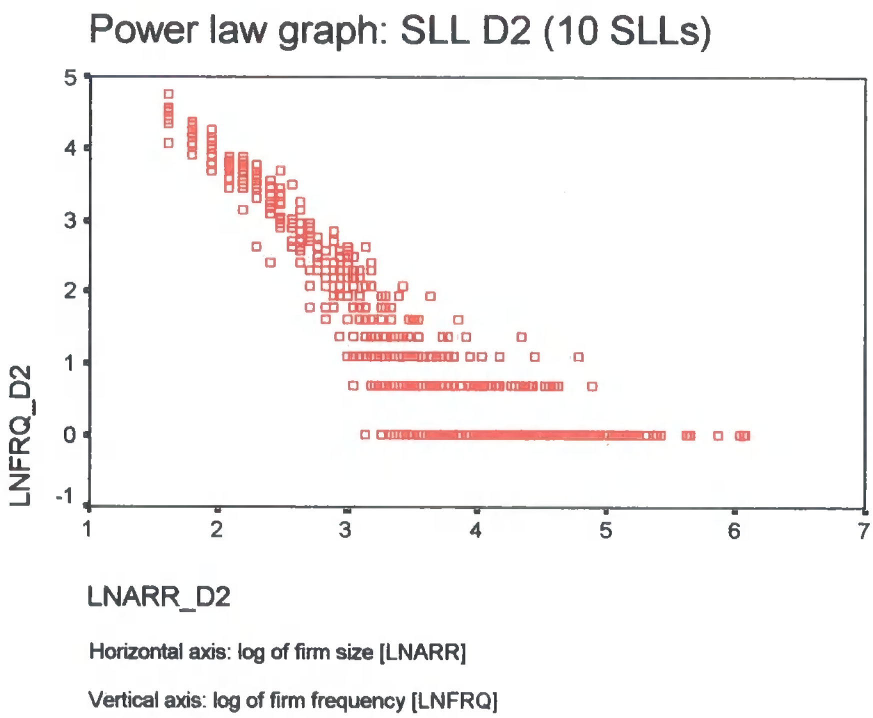


Figure 39: Scatter graph of power law for D2 (10 SLLs)

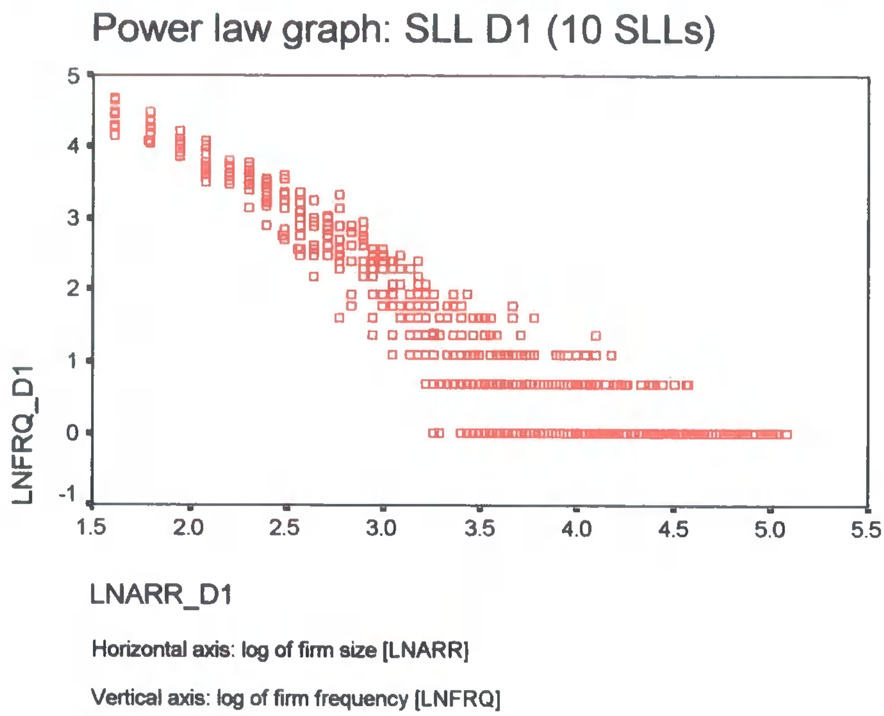


Figure 40: Scatter graph of power law for D1 (10 SLLs)

As it is in the case of the single clusters, also the aggregates of ten SLLs per category follow a power law. Two implications follow from these results: the first broad implication is that the networks (and the aggregation of networks in categories) under analysis do not belong to the random type. Contrary to intuition, they do not show a typical scale, which in this case, should be a preferential size, with all the other values rapidly decaying on either side of the mean. In other words, the first result indicates that local industrial agglomerations do not exhibit the properties of random networks[515], but instead the scale-free property of power laws. Second, as discussed in the second paragraph of section 4.2, closeness to a power law can be further inquired by means of a regression analysis in order to look for differences in the degree of fitting of the distributions to a power law. This especially important in the case of industrial agglomerations as in each type a different degree of self-organisation is assumed to be present.

[515] See footnote 256

The SLLs I have selected for this research are shown in Table 9[516].

A note of caution however concerns the truncated population used for this research (lack of firms with less than 5 employees). It is possible that by including the rest of the population a distribution closer to a random network could still emerge.

The results of the regression analysis for the four categories are summarily shown in Table 8. This more rigorous analysis shows significative differences between the four SLL classes. The meanings of symbols are as follows: N_data indicates the total available number of frequency-size records per category. R^2 represents the regression coefficient, B stands for the slope of the regression equation and B StErr gives an indication of the spread of the value of the distribution. Cut-off values indicate the average number of employees at the 99% cut-off point.

	N_data[517]	R^2	B	B StErr	Cut-off
A2	560	0.73	-1.10	0.03	387
A1	560	0.74	-1.19	0.03	302
D2	560	0.79	-1.31	0.03	227
D1	560	0.85	-1.54	0.03	139

Table 8: Main statistical indicators of power law analysis

The results indicate that the regression coefficient and the slope of the distributions seem to follow the progression of categories from less to more clusterised agglomerations (that is, from A2 to D1). It is part of the definition of the categories that the shift of industrial characteristics from A2 to D1 reflects the progressive weight of micro and small enterprises over total employment and the increasing specialisation of the SLL's firms around some macro-sectors. As these two are essential features of the Marshallian 'district' (cluster), it is legitimate to claim that the cluster 'content' increases from A2 to D1[518]. As R^2 is "*a measure of the closeness of fit of the regression equation*

[516] see methodology (section 4.2.2.2) for more details regarding the parameters used for the selection

[517] for statistical comparability the number of entries per categories has been equalled to 560, which is the lowest among the four categories: A2=560; A1=652; D2=639; D1=609. N_data differs across categories as it depends on the span of the array, that is on the number of different sizes charaterising the populations.

[518] See section 4.1.2 for more details

to the sample data"[519], I can conclude that the cluster content of the SLLs correlates with the power law and therefore that the proposition seems to be validated by the data.

SLLs more likely to include industrial clusters are closer to a power law than the other types. As the cluster 'content' correlates positively with self-organisation, I can legitimately conclude that a power law distribution is an indicator of self-organisation in action.

In short what we observe from the results shown above is:

1. A correlation between the regression coefficient and the "cluster content" of the SLL
2. A correlation between the slope coefficient and the "cluster content" of SLL
3. A correlation between cut-off point (that is, number of employees of firms at cut-off point 99%) and "cluster content".
4. It seems that the four categories can be regrouped into two groups: A2, A1, D2 on one side and D1 on the other side. The difference of D1 with any of the other categories exceeds the internal differences between themselves.

These points will be discussed in the following chapter.

---

[519] (Daniel and Terrel 1989) p.480

SLL code (ISTAT)	Name SLL	Number of councils	Number of local units	Total number of employees	Type of SLL
27	ALBA	50	7973	35290	A1
44	ALESSANDRIA	32	12543	58328	A1
40	ASTI	50	9681	40758	A1
196	BELLUNO	16	6213	30842	A1
33	CUNEO	51	11888	52732	A1
123	PAVIA	52	12943	66313	A1
383	PERUGIA	7	15493	76516	A1
236	PIACENZA	19	13837	61974	A1
349	PONTEDERA	15	9236	38681	A1
74	SESTO CALENDE	38	9151	42941	A1
142	BOLZANO	15	12645	69932	A2
273	CESENA	10	15056	58117	A2
415	FROSINONE	35	15282	72344	A2
135	MANTOVA	14	11547	55747	A2
231	MONFALCONE	28	10341	42309	A2
22	NOVARA	25	11580	63857	A2
495	PESCARA	10	17989	81967	A2
347	PISA	5	12654	59468	A2
389	TERNI	15	10446	52164	A2
175	TRENTO	31	12157	66461	A2
213	CITTADELLA	17	9472	44101	D1
188	BASSANO DEL GRAPPA	20	11580	54679	D1
250	*CARPI*	3	8617	39261	D1
90	DESIO	37	34621	168164	D1
72	GALLARATE	21	14594	81745	D1
110	LUMEZZANE	14	5687	29487	D1
204	*MONTEBELLUNA*	16	9000	43738	D1
114	PALAZZOLO SULL'OGLIO	11	6065	30780	D1
339	*PRATO*	9	28152	101700	D1
350	SANTA CROCE SULL'ARNO	6	8792	38255	D1
313	*ASCOLI PICENO*	15	7630	35951	D2
11	*BIELLA*	44	10568	49542	D2
202	CASTELFRANCO VENETO	12	9639	43938	D2
130	CREMA	34	7252	32837	D2
274	FORLI'	6	13188	58569	D2
253	MODENA	9	21524	113368	D2
322	SAN BENEDETTO DEL T.	11	9860	38490	D2
257	*SASSUOLO*	5	9213	53559	D2
194	*VICENZA*	23	19240	102193	D2
126	VIGEVANO	28	14640	78081	D2

Table 9: Main indicators of selected SLLs

5.3 DIVERSITY RESULTS

5.3.1 INTRODUCTORY REMARKS

This section presents the evidence supporting the proposition on diversity:

Self-organising systems maximise systemic diversity (product of diversity times connectivity) at a rate higher than aggregates (and hierarchical systems).

I have introduced in section 4.3.1 the definition of diversity I have adopted for this research work and introduced the algorithm used to measure diversity and connectivity in industrial agglomerations. The theory and rationale behind the propositions is presented in chapter 3 (theoretical framework). I will briefly recall the important variables necessary for the measurement of diversity. Generic diversity (called simply diversity) is the product of variety (the amount of measurable industrial species), times balance (the density of organisations per species), times disparity (the distance between the categories used to describe the species). Given an industrial agglomeration (SLL) and a number of firms per agglomeration, the measurement of diversity is taken as the relative diversity between any two pair of firms summed[520] over all pairs. Distance between species is taken as the algorithmic distance between the digits of SIC codes. Systemic diversity is instead the product taken per each pair of firms of their diversity times connectivity, again summed over all pairs. Connectivity is assumed to be proportional to the financial value of input-output transactions between industrial species. All results are based on a sample consisting of 600 hundred randomly extracted firms per SLL. The SLL constitutes the unit of analysis of this research. The results presented in the next section show how the various features affecting diversity change across the four SLL categories. The details of the calculations are presented in section 4.3.4.

I will in the following section present the evidence starting with the measurement of variety.

5.3.2 MAIN RESULTS

[520] as diversity is an additive measure and clusters have different populations, a random selection of 600 firms has been used to equalise clusters' populations across all categories

Variety

Variety is measured counting the number of 3 digits SIC codes[521] in each SLL and taking the average over 10 SLLs in each category. It reflects the spread of specialisations within the categories. The results are shown below in Figure 41. It is interesting to notice that D1 variety is the lowest and significantly different than in the other cases.

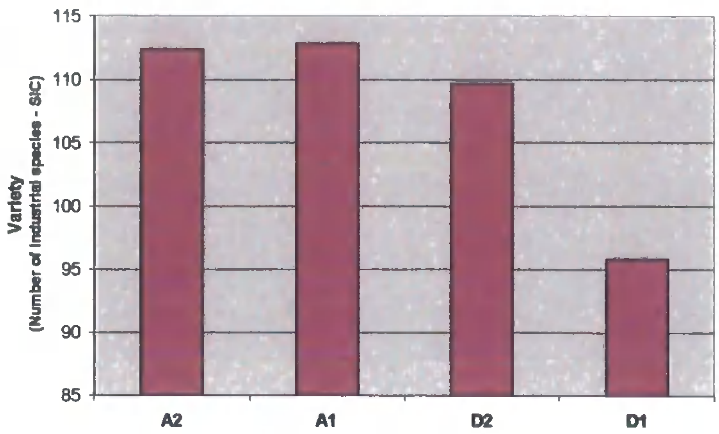


Figure 41: Variety – average number of industrial species (SIC codes) per category

Figure 42 presents the ratio[522] of total number of firms per category divided by variety (the spread of specialisations). The ratio provides an indication of the overall dispersion of firms across the niches (or specialisations). Figure 42 confirms the expectation that specialisation increases from A2 to D1 and therefore that the cluster mechanism encourages aggregations of firms around a more limited sets of specialised niches.

---

[521] the choice of 3 digits is forced by the fact that input-output (I-O) tables are available only at 3 digits resolution. As variety is a multiplicative factor in the calculation of systemic diversity, which rely on I-O tables to calculate connectivity, the adoption of 3 digits turns out to be necessary for the correct implementation of the algorithm

[522] the ratio is obtained by dividing the total number of firms by the total number of SIC in the unfiltered population.

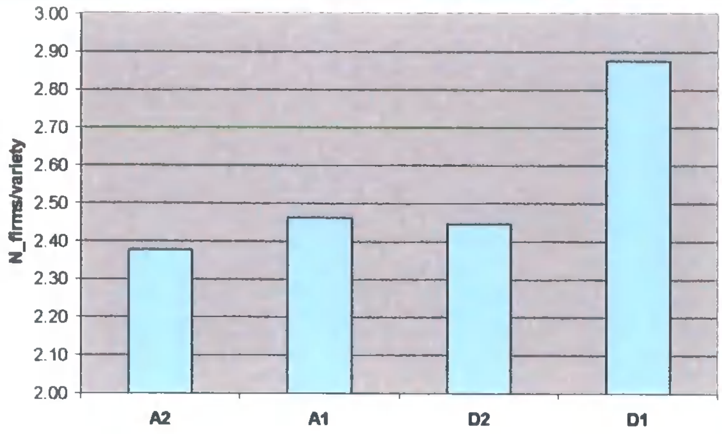


Figure 42: Specialisation index - average number of firms per niche in each category

Distance

The value of the distance (otherwise defined disparity) represents the sum over all pairs of the relative distances between any pairs of firms in each category. The measure of distance provides an indication of the relatedness of industrial activities in each category. Figure 43 shows that A2, A1, and D2 exhibit a higher dispersion of activities than the more compact D1. This implies that true industrial clusters (D1) are characterised on average by patterns of more closely inter-related activities. In other words, the coherence among D1's activities yields a lower distance.

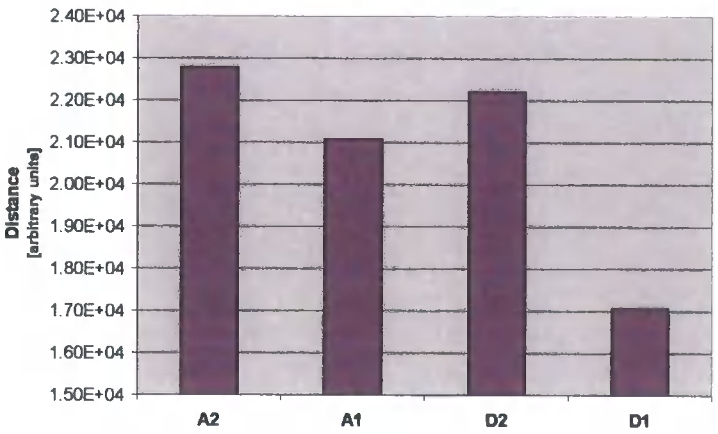


Figure 43: Distance – distance calculated as sum over all pairs of firms of the algorithmic difference between digits in SIC codes (in each category) [arbitrary units]

Diversity

Figure 44 shows the measure of diversity. Diversity provides an effective measure to describe the spread of industrial species in a cluster. It constitutes a more complete measurement of diversity than variety alone. It is interesting to notice from this point of view that D1 shows the lowest diversity of all categories.

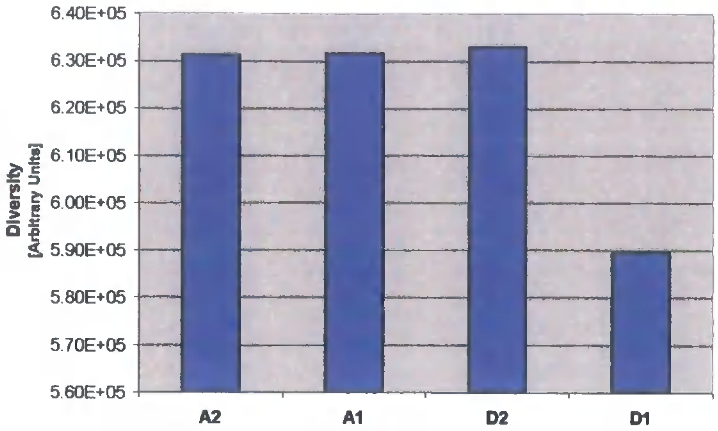


Figure 44: <u>Diversity</u> – generic diversity or sum over all pairs of distance and variety in each category [arbitrary units]

Systemic diversity

The measure of diversity is not concerned with the linkages between firms. It is possible that firms sharing the same territory have no connection with one another or on the contrary that they belong to the same value chain. To understand more as to how diversity is related to interactions among co-located organisations, diversity has to been seen not as a list of static differences, but rather as a map of complementarities among agents. In order to do so, each pair of firms is multiplied by a weight factor, proportional to the average financial transaction between the two industrial species (SIC codes) to which the firms belong. The weight factors, defined as connectivity factors, are obtained by financial input-output tables. Systemic diversity represents the direct measure of the diversity-connectivity product, which is the object of the proposition to be tested. Figure 45 confirms that agglomerations with the highest cluster 'content' show the highest systemic diversity, thereby confirming the proposition.

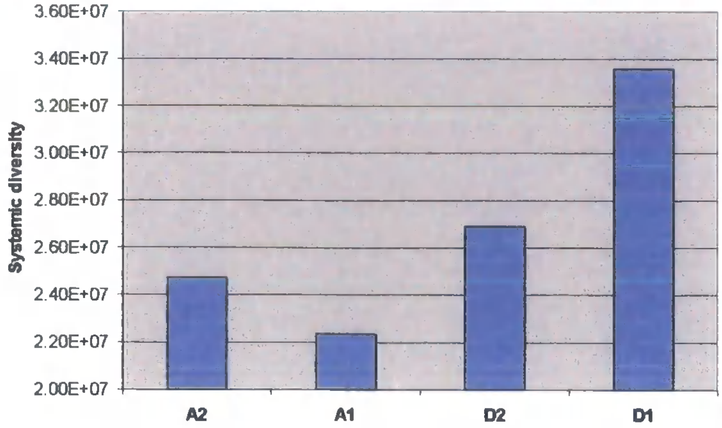


Figure 45: Systemic diversity – product of diversity times connectivity [arbitrary units]

5.3.3 DIVERSITY AND CONNECTIVITY IN THE MOST IMPORTANT INDUSTRIAL CLUSTER IN ITALY: PRATO

The approach presented above allows the calculation of the basic features of diversity across different types of industrial agglomerations. The approach has confirmed that a general framework based on a dynamic interpretation of diversity is able to measure (in quantitative terms) the cluster effect in industrial agglomerations. But what happens if instead of relying on the Cannari & Signorini classification of SLLs, we take a different approach and compare firms that belong to the core activities of a cluster with firms that do not within the same agglomeration? The driving question becomes whether the distinction between the dynamic typical of systems and the dynamic typical of aggregates runs through single agglomerations. To carry out this analysis I have selected a single SLL. The choice has fallen onto Prato for the following reasons: first, Prato is one of the most famous industrial clusters[523]; second, it offers a large population of firms for statistical analysis; third, it is easier in Prato to draw a distinction between firms that belong to the core textile cluster and firms that do not, due to the extensive literature existing on the district.

In this case the distinction between different types of agglomeration dynamics is internal to the cluster, between activities (always defined by SIC codes) internal to the cluster's core value chain and generic activities. Drawing this distinction is partly a matter of judgement. There is no doubt that in a cluster strongly dominated by the textile

---

[523] (Best 1990; Porter 1990), (Piore and Sabel 1984; Becattini 1997)

sector such as Prato, CAD/CAM specialised firms, design and marketing agencies (just to name a few) are shaped by the nature of the cluster's products and way of organising business. However, the distinction becomes fuzzier when the transportation activities are taken into consideration. Some of them are specialised to serve the cluster's specialised products, others are not. The distinction between core and non-core activities is based on the breakdown of three digits SIC codes between core and non-core industrial activities (for a total of 222 including services) and is the work of Andrea Balestri[524], chief of the "*Centro Studi*"[525] of the "*Unione Industriale Pratese*". As for the previous analyses, a random selection of 600 firms for each category (core and non-core) has been taken. The results are shown below.

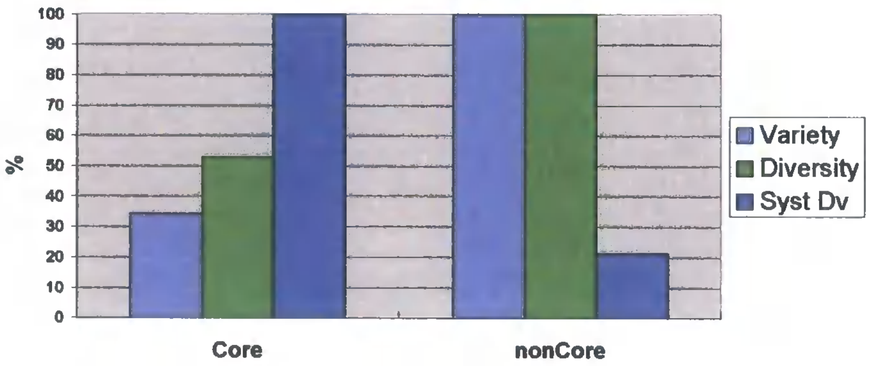


Figure 46: Variety, diversity and systemic diversity in Prato [percentage values]

	Core	nonCore
Variety/N_firms	31	90
Diversity/N_firms	15	29
Connectivity/N_firms	54	528
Systemic diversity/N_firms	5663	1202

Table 10: main data for variety, diversity and systemic diversity in Prato (all data are normalised by the number of firms (600)

What does it emerge from this analysis? The trends seen at the national level between the four categories of industrial agglomerations are confirmed by the Prato case. Even more, the differences between cluster and non-cluster activities (reflected in the

[524] I interviewed Andrea Balestri in June 2001 in Prato

[525] Study Centre of the Confederation of the Prato Industries.

core/non-core distinction) are amplified. To compress several variables into the same graph, the variables shown in Figure 46 are expressed in percentages. The real numbers are shown in Table 10. The point I want to stress is that the converging results between the multi SLLs case and the Prato case are obtained in two very different contexts: the first class of results is based on the SLL as the unit of analysis and describes the properties of diversity in SLLs across macro-classes of aggregations (given by the Cannari & Signorini classification); the second class adopts an *ad hoc* grouping of firms (divided in core cluster and non core cluster activities) within one specific SLL as the unit of analysis and describes the properties of diversity in a more micro environment. The fact that the nature of the results in the two cases is the same constitutes a strong validation of the proposition. A full discussion of the empirical results is devolved to the following chapter.

Chapter 6: Discussion of results

6.1 POWER LAW DISCUSSION

The evidence presented in chapter 5 shows that the distribution of firms' sizes is well approximated by a power law, irrespective of the agglomeration type. As a power law distribution is an indication of an interconnected dynamic (among the constituent part of the agglomeration under analysis), this result indicates that all four categories of SLLs are closer to systems than to aggregates. In itself this result is not totally unexpected. The structure of Italian industry, dominated as it is[526] by small and micro enterprises, linked along complex supply chains and cooperative linkages, shows, almost inevitably, systemic features. From this point of view the emergence of systemic features across all four categories of agglomerations (from non specialised industrial areas to industrial clusters, that is from A2 to D1[527]) is hardly a surprise. Complex economic activities can not take place in a vacuum. However, a visual analysis of power laws is not an appropriate method to resolve qualitative differences between structurally different systems, that is, differences that concern the mix between self-organising, random and hierarchical properties in social systems. The fact that all the types of industrial agglomerations show elements of systems does not imply that they are characterised by the same mix of properties mentioned above. In what follows I present a more sophisticated approach to identify the differences between the categories of agglomerations.

6.1.1 REGRESSION COEFFICIENT

A regression analysis gives a measure of the closeness of fit of the regression equation to the sample data, thereby providing a test for proposition 1. The results reported in Table 8 and in Figure 47 (below) indicate that the fundamental proposition seems to be confirmed by data[528].

Two questions arise:

---

[526] See Table 2 in chapter 4

[527] see 4.1.3

1) What is the reason for this distribution?

2) What are the forces that bring the power situation into being?

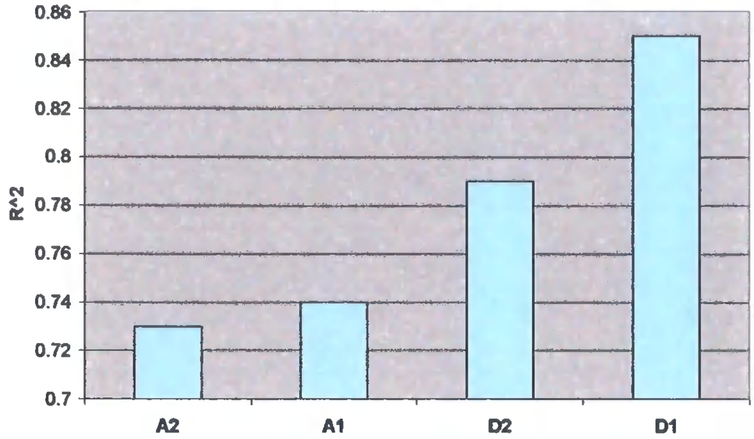


Figure 47: Regression coefficient for the four categories (10 SLLs per category, cut-off 99%)

6.1.1.1 Self-organisation and power law

The first question leads directly to the question of what an industrial cluster is. We have seen that a power law is an indicator of a complex system in the dynamic state known in literature as "*the edge of chaos*", in which chaotic dynamics coexist with islands of order and stability. The *edge of chaos* is reflected in the fundamental feature that sets complex systems apart from externally-engineered systems, or systems controlled by deterministic laws, that is, their capability for self-organisation, which consists in the optimisation of the mix between order and chaos in order to achieve at the same time homeostasis and innovation[529]. Does this description apply to industrial clusters?

It does because the structure of industrial clusters is one of adaptive networks in which the multiple feedback channels between the constitutive firms generate a coevolutionary dynamic and consequently a self-organising behaviour. Clusters can be seen as a form where internally coherent and highly connected local pockets (formed around *ad hoc* projects) can coexists and interact with other pockets. The intra-pocket

---

[528] In section 5.2.1 I have also shown that the differences between the four categories are statistically significant (less than 5%) at least for the SLL of type D1

[529] the autonomous generation of the attractor dynamic is instead due to symmetry breaking and bifurcations; see section 2.3

links ensure stability whereas the weaker inter-pockets links ensures that pockets maintain their identity whilst, at the same time, making possible the innovation-generating interaction of the diverse pockets. We have that the *"diversity of organisations"* creates the rules for the *"organisation of diversity"*[530].

If we consider industrial clusters as an intermediate organisational form between on the one hand simple aggregates of firms without any stable system of relationships and on the other hand integrated companies (see Figure 48[531]), it is easy to see that self-organisation emerges only in the middle and is absent in the extremes of the spectrum. Industrial aggregations are quasi-random agglomerations in which individual agents are free to pursue their fitness limited only by generic environmental constraints[532]. Much of their activity is dissipated randomly without the construction of structures of order, which operates at the system level, like molecules of the Bènard fluid in the chaotic state. They self-manage, but don't self-organise in the meaning that their interactions do not generate structures of order that transcend the single agent's actions. The embodiment of this type of production structure that Fujita defines as *backyard capitalism*[533] is represented by semi-autarchic entities, which produce much of what is needed for their self-consumption. Moreover the small dimension and limited connectivity characterising this type of organisations limit self-organisation and the emergence of higher level structures of order. On the other hand, in the integrated company case, especially in the Fordist type, self-organisation is ruled out by the hierarchical approach of command and control and by the minute subdivisions of tasks in always smaller tasks, hierarchically controlled by the brain of the organisation[534]. In the middle the network structure allows a) the freedom of the chaotic aggregation (although shaped by a set of constraints set up by the multiple relationships and feedback loop system) and b) the coordination and structural properties of the integrated case. An example of a collective property is the industrial cluster's extreme agility in product innovation, based on the capability of

---

[530] the expression is by (Grabher and Stark 1997)

[531] my figure

[532] Constraints can under certain circumstances build complex systems. For a theory of how this happens see (Juarrero 1999). Kauffman instead discusses a mirror-like approach: how constraints generate more specific ways of utilising energy (Kauffman 2000)

[533] (Fujita, Venables et al. 1999) p.2

[534] I am not claiming that self-organisation does never happen in vertically integrated organisations, only that the command and control style create a 'hostile environment' for self-organisation.

quickly reconfiguring the cluster's nodes (in the introduction and in section 3.3.1 this point is developed at length).

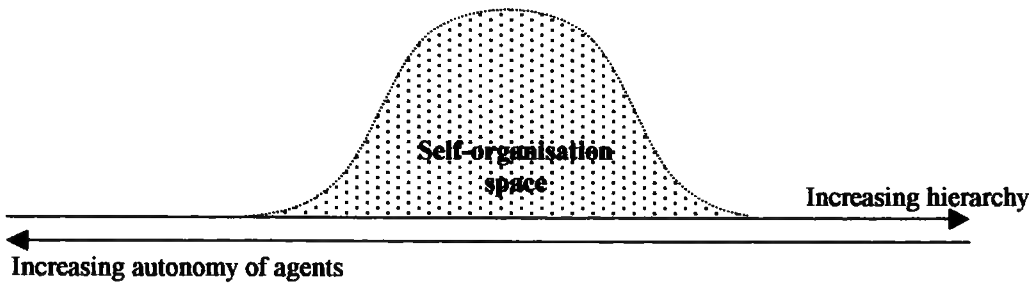


Figure 48: Self-organisation spectrum

The theoretical consideration that industrial clusters are complex systems on the *edge of chaos* is confirmed by the positive correlation between power law and 'cluster content'[535] of industrial agglomerations (SLLs). SLLs closer to the aggregate structure (A1 and A2) are further away in terms of regression coefficient from a power law, in line with the proposition that aggregates show elements of random networks. The fact that SLLs with higher cluster content (D1) correlate better with a power law confirms that the dynamic that leads to the dimension of the nodes is not random, but instead self-organising.

However, the complete validation of the proposition regarding the correlation between self-organisation, clusters and power laws requires a similar measurement for the case of the integrated organisations. This measurement poses complex methodological problems and requires access to sources of a completely different type from the ones used in this research. This measure will not be attempted in this research.

In conclusion, the comparison between different types of industrial agglomerations using power law theory indicates that self-organisation and power law are correlated.

6.1.1.2 Principles of power law

But what are the forces that cause power law distributions to emerge? To answer this question we need to go back to Simon's model (see section 3.2.1) or to Barabasi's

[535] By cluster content I mean the probability of industrial agglomerations (as defined in the Cannari & Signorini classification, see section 4.1.3) of including a Marshallian cluster.

reformulation[536] of the same principles. Simon showed that when lumps (let's say entrepreneurs) of a certain size colonise new nodes in a network or add to the manpower of existing nodes, under certain assumptions a power law is found. Is the situation in industrial clusters very different?

Behind Simon's model there are three assumption (also described earlier as Krugman's assumptions), which are: a) the phenomenon is dynamic, b) the growth is random and c) scale-independent.

If these three conditions are verified in the case of industrial clusters but not in the other two cases (aggregates and integrated organisation), the emergence of a power law will become less mysterious.

Integrated organisation[537]

The first condition is verified when there is growth (either positive or negative). However, a power law will emerge and persist only in systems characterised by persistent growth. This may happen for different reasons, for instance turbulence in the environment or persistence in the generation of internally generated innovations. In both case this requires the presence of dissipative conditions, which generate apparently unstable systems that are kept together by the constant renewal of their organisational complexity[538]. As the survival of the integrated organisation is guaranteed by (among other things) the implementation of efficiency seeking strategies, which are mostly compatible with stability of markets and technologies, the condition of permanent change is not automatically verified. During periods of expansion or contraction, the reaction of the system to external environmental changes (adaptation) is planned and filtered in a top-down fashion from the executive level. Random change of the organisation's parts is actively discouraged. The system behaves as a close system with carefully controlled channels of interaction with the outside world.

In conclusion, as the rational for the integrated organisation is to exploit size in order to internalise the market dynamics that generate instabilities, with the ultimate goal to

---

[536] Barabasi does not seem to be aware that he is moving on Simon's steps or at least this awareness does not emerge from his book (Barabasi 2002)

[537] this section refers to the Fordist organisation as described for instance in (Piore and Sabel 1984)

[538] the concept that dis-equilibrium situations can drive a further increase in dis-equilibrium is modelled by (Allen 1997) p.147

create monopolistic conditions[539], the final effect is to deter change, both at the environmental and internal level. The integrated organisational form is based on stability-pursuing strategies and is therefore a successful strategy in presence of long-term incremental change. The other two conditions, random and scale-free growth do not hold either due to the centralised decision-making arrangement.

Aggregate

With regard to the first condition, the aggregate can be either stable or under change. As the aggregate is a collection of parts, the question of stability of the aggregate really depends on the stability of the constitutive parts. Reductionism applies here. At least at the aggregate level, there is no inherent necessity of permanent change in order to ensure survival. The aggregate can be stable without that causing a threat to its survival. Again the first condition is not automatically verified. The second condition is verified. There is no reason why growth should happen in a part rather than in another of the aggregate. The third case is more interesting: if one of the firms in the aggregate starts getting ahead of the others, it is to be expected that its potential for attraction of more lumps will increase with its size, essentially because of economies of scale. As economies of scale are based on increasing returns mechanisms, it is plausible that integrated organisations will enjoy higher economies of scale than small isolated companies[540]. Beyond a certain dimensional point dis-economies of scale will also set in and limit further expansion[541]. Usually the balance between economies and dis-economies of scale remains positive for a large dimensional range and tips in the opposite direction for large vertically integrated organisations. Aggregates are unstable to small perturbations[542] that make the distribution of firms collapse toward one dominated by few small firms. The critical point here is that economies of scale are inherently non linear and, if not limited by diseconomies of scale or other types of centrifugal forces, would inevitably generate a

---

[539] (Piore and Sabel 1984)

[540] several increasing returns mechanisms are reported in (Arthur; Arthur 1996). For the very relevant discussion of collapse of spatial homogenous distribution of producers into few hot spots see (Fujita, Venables et al. 1999) and (Krugman 1996)

[541] see Arthur on Marshallian economics (Arthur 1996)

[542] This mechanism is beautifully explained by (Schelling 1978)

'*winner takes all*' type of situation[543]. This distribution would not obey a power law because of this "increasing returns" type of situation.

There is a further point: the aggregate behaves as a (very) open agglomerate. The driver of growth can be as much endogenous as exogenous. Reference to Figure 27 in chapter 4 shows that the ratio of local units to firms is the highest for industrial areas of type A2 and it decreases going towards industrial clusters (D1). This indicates that in non-cluster areas the percentage of firms that are part of larger groups (with several local units), and therefore more likely to have the controlling unit outside of the SLL, is higher. Growth in non-cluster areas is more sensitive to decisions taken outside the area by large firms and therefore non random.

This discussion confirms that Simon's conditions do not apply to the aggregate case. This is confirmed by the statistical results that indicate that industrial agglomerations, which are more likely to include features of aggregates, show the lowest correlation with a power law distribution.

Industrial clusters

The pure form of an industrial cluster is necessarily dynamics. Why? A cluster is after all a set of relationships and competencies - partially residing in nodes and partially distributed across the whole cluster. The cluster as a meta-organisation relies on economies that are external to the individual organisations but internal to the cluster. The critical point here concerns the appropriability of the returns of those economies. Because they are distributed across dynamic networks, which form around specific projects and products/services and are therefore temporary, the returns are appropriated by the network itself, through which they eventually flow back into the whole community. For instance, a micro-innovation, which contributes to the success of a network within the cluster, spreads rapidly within the cluster[544], even though its knowledge is largely tacit and based on *learning by doing*. In the medium term, the returns of the innovation are appropriated by the cluster, as the diffusion of the innovation within the cluster is fast whilst outside is slower due to the embeddedness of knowledge and routines within the cluster. However, although the distributed nature of

---

[543] (Arthur, Durlauf et al. 1997) (Cusumano, Mylonadis et al. 1997) (David 1985) (Gladwell 2002)

[544] (Rullani 2001) has described innovations happening in clusters as being inclusive toward the internal environment, that is subject to rapid diffusion, and exclusive toward the external environment.

the cluster form is subject to the risk of implosion due to the centripetal action of economies of scale and therefore to the loss of the collective appropriability of innovations, this rarely happens. The argument is as follows. If some firms get control of their partners and transform them into suppliers along the lines of traditional supply chains, and if they internalise services and production that are currently externally provided, then the external economies would become internal economies and the cluster would be transformed into a more traditional aggregate of large integrated companies. This transformation would probably in the short term be beneficial to the controlling company but detrimental to the survival of the cluster. In the medium term the transformation to vertical integration would isolate the company from the flux of knowledge and innovation that takes place at the heart of the cluster and therefore put the company at a disadvantage[545]. The external economies deriving from the rapid cluster appropriability of individual firms' (or networks') results would simply disappear. Unless the market evolves toward stability, where the economies of scale argument offsets the advantages of flexibility and agility, the evolution toward the integrated model is detrimental for competition in turbulent markets. Silicon Valley's large firms tried to compete with the Japanese companies in the semiconductor markets in the eighties by moving toward vertical integration[546]. The attempt was a failure because of the incompatibility of the integrated model with a highly innovative environment. In fact, the Valley reverted back to the successful cluster model with the advent of the multimedia and Internet wave of the nineties. If previously clusterised organisations integrate vertically, the cluster simply disappear. In fact, clusters are meta-organisational forms that thrive on continuous exploration and exploitation of market niches, that open and close at the boundary of mass market productions. It is a maximally diversified structure, able a) to maximise its potential of attracting buyers thanks to its unmatched variety of products and flexibility of offers and b) to scan the environment searching for new un-exploited (and co-created) niches, thanks to the variety of its detection systems. In order to identify, occupy and exploit these highly profitable niches, clusters have to retain the structure based on a web of self-organised network. In fact, niches exploitation requires

---

[545] I have described in the introduction (see section 1.8) the transition from a cluster to a traditional structure due to similar mechanisms.

[546] In the 80s (Saxenian 1994) many firms in Silicon Valley acquired traits of traditional integrated companies in response to an environmental change away from leading edge innovation and toward mass marketing manufacturing. That was the period of the Japanese dominion in the memory industry and of the emergence of a clear winner-takes-all standard in the PC industry (IBM/Microsoft/Intel standard)

a continuous reconfiguration of specialised modules that are packaged together and disbanded when the niche potential is over (see section 3.3.1). The mechanism of reconfiguration is key to the collective appropriability of returns.

For the reasons seen above, there is then an irreducible dynamics at the heart of clusters. Stable markets are incompatible with the cluster form as this organisational form is utterly unable to exploit the economies of scale that are necessary to survive in mass-markets. Clusters can only survive at the edge of chaos because they are far-from equilibrium meta-organisations. In conclusion the first condition is always verified in clusters.

Let's examine the second assumption, random growth). In networks not of the hub-and-spoke type there is no centre and therefore no reason for asymmetric growth. Also the relative absence of large firms implies that there are very few players whose decision-making can significantly affect the evolution of the cluster. Growth is distributed and random. The second condition holds.

The third assumption (growth is linearly proportional to the size of the nodes – or to the number of links) is really about the reasons why economies of scale fail to produce an increasing returns mechanism of single nodes growth. The first reason, having to do with external economies and appropriability of rents, has been discussed earlier. The second instead regards the mechanism of phase specialisation. Clusters, especially in the Italian version, are mechanisms in which deepening specialisation takes place via organisational unbundling. Clusters are constantly exploring the frontier of indivisibility of tasks[547] in the attempt to maximise their internal diversity. This is useful to increase flexibility and avoid internal me-too competition (Shumpeterian competition – see section 3.3.1.3). When a possibility for further phase specialisation (indivisibility turns out to be divisible) surfaces due to technical progress, scientific discovery, diffusion of innovation or changes in customers' demand, then organisations respond by de-merging, spin-offing or outsourcing elements of previously integrated phases. This causes the emergence of start-ups dedicated to the new phases. It is the constant emergence from the top (spin-off) and from the bottom (start-ups) of organisations, more specialised than the environmental average, which limits the emergence of economies of scale internal to the organisations. Cluster's organisations are kept in balance by two opposite forces: the

---

[547] (Becattini 1990)

first is the tendency toward internal growth in order to minimise transaction and fixed costs by exploiting economies of scale; the second is represented by the tendency toward fragmentation of existing organisations and formation of new ones. These two forces act synchronously on organisations. It could be argued that the type of balance between economies and diseconomies of scale is similar. This is not so, as diseconomies of scale[548] become active after the reaching of a set of thresholds, thereby setting different ranges for the action of the two forces. The phase specialisation mechanism instead operates as an adaptation to turbulence in markets and technologies and its action is therefore not restricted within particular ranges.

It is clear therefore that Simon's model applies better to industrial clusters than to the other two organisational forms. The forces that drive the emergence of power law in industrial clusters are the networked modular structure of clusters and the incidence of phase specialisation (through the bubbling up of more specialised organisations) in the context of persistent external economies. The process of growth is necessary, random and size-independent, and takes place via an exploration of the *adjacent possible* in terms of innovation and search for the indivisibility frontier of production and organisational techniques, actuated via a constant demerging of specialisations.

6.1.2 CUT-OFF POINT

The analysis of the slope coefficient reveals that D1 districts seem to differ substantially from the other three groupings.

The slope coefficient (see Table 8 at page 215) is an indication of the steepness of the distribution. If the slope was exactly 1, then for each company of size n, there would be 2 companies of size n/2, 3 companies of size n/3 and so on, according to the law

$$F \approx S^{\alpha}$$

F stands for the number of companies with size S or bigger and alpha represents the slope.

The results indicate that the ratio of small to large companies is larger in a district of type 1. The reasons suggested above, related to increasing phase specialisation achieved

[548] (Arthur 1996)

by fragmentation and spin-off, limit the growth of larger companies and prevent vertical integration. This also causes a further consequence related to cut-off point.

The history of Prato[549] shows that the birth of the industrial cluster (roughly between the end of the 50s and the first half of the 60s) was accompanied by a radical decimation of companies larger than 500 employees to the point that no one survived. This phenomenon is common to many industrial clusters. The organisational dynamics introduced in the first three chapters - having to do with the networked modular structure of industrial clusters, phase specialisation and recombination of modules around *ad hoc* projects – explain the trend toward smaller firms. My analysis confirms that industrial agglomerations closer to the cluster form include mechanisms that put large firms at a disadvantage. The analysis shows that the maximum average size of firms depends on the type of agglomeration. To calculate this point I have analysed the 99% cut-off point of the cumulative frequency distribution. This represents the size of the company (in terms of number of employees) that is 1% close to the largest company in the agglomeration. Taking the 99% cut-off point instead of simply the largest company reduces the uncertainty connected to casual events, for instance the presence of a very large company located there for reasons other than endogenous. Also the 99% cut-off reduces the effect of the error in the database highlighted in section 4.2.2.1, which concerns the inclusion in the local agglomeration of firms in their entirety and not in terms of their local units only.

The results, shown in Figure 49, indicate that the average cut-off point for clusters of type D1 is much smaller than for the other categories (see Table 11). The figure also shows an almost linear decrease[550] of the cut-off point between the four categories. The standard deviation of the cut-off points relative to the four categories however shows a different picture. The regularity of the average is substituted by the emergence of two groups: on the one hand, there are the categories A2, A1 and D2 that show a very high standard deviation and, on the other hand, the very compact category D1. The discussion of this point is postponed till the next section.

[549] (Becattini 1997)

[550] the trend is not surprising in itself: in fact one of the criteria that define the SLL classification is the percentage of small firms as compared to the national average (see section 4.1.3)

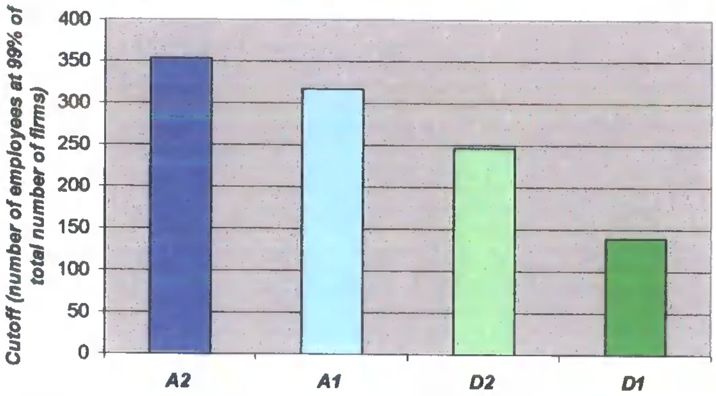


Figure 49: Cut-off point for the four categories

Category	Average cut-off value	Standard deviation
A2	387	108
A1	302	127
D2	227	198
D1	139	35

Table 11: Average and standard deviation of cut-off points for the four categories

Alternatively, one may suggest that the predominance of SMEs in clusters is not due to inherent dynamic phenomena specific to clusters (spin-offs, disintegration, etc) but is simply the result of the fact that clusters of type 1 are dominated by local enterprises with local markets. These, by definition, consist predominantly of medium and small enterprises. The other areas instead are characterised by the presence of bigger regional, national and international companies. According to this view, the relevant distinction is not between clusters and non-clusters but instead between local and more 'international' areas. This explanation captures a part of the truth. Clusters are dominated by local enterprises with no or very little presence of exogenous industry (as Figure 27 in chapter 4 shows), but, as the example of Prato (and the story told in the introduction) demonstrates the cluster form is achieved via a process of disintegration of large companies and not via a permanent lack of these. Also, although the origin of companies and (generally speaking) business dynamics are local (to the point that it would be better to use the word idiosyncratic), clusters are strongly export orientated, much more than the other groups in our taxonomy[551]. In conclusion the 'localness' of cluster does not

[551] (Bronzini 2000)

represent an alternative explanation, but is a necessary complementing feature of the definition of clusters.

This result regarding the cut-off point confirms the evidence provided by the single case of Prato and gives further support to the phase specialisation mechanism mentioned above.

6.1.3 Regrouping of Classes

As suggested in section 4.1.3.2, the Cannari & Signorini's refinement of the dichotomic (cluster versus non-cluster) division of SLLs, allows a more fine-grained exploration of the cluster properties as opposed to more traditional forms of classifications. The surprising finding is that the agglomerations of type D2 traditionally considered as clusters are instead closer to non-cluster agglomerations of type A1 and A2. This finding raises some important issues:

First, the methodologies used to identify industrial agglomerations are usually based[552] on some econometric or demographic criteria, such as index of specialisation, incidence of SMEs on the total of firm population, etc., which are matched with geographic data. These criteria are very effective but coarse-grained, thereby overlooking some of the finer issues. In general they tend to measure either properties at the nodes level (firm), or at the aggregated level (for instance, SLL's aggregated income). However, complex systems link nodal properties to systemic properties through emergent coupling mechanisms. Interlevel dynamics characterised by the simultaneous presence of bottom-up and top-down mechanisms, by which the concerted actions of agents generates the systemic coherence of the whole and the latter defines the context and the rules of engagement for agents, create a methodological problem that in a reductionist world would not exist. In fact, measurements at one level can not be assumed to describe the upper more inclusive level, as the inter-level summation inference loses meaning. In this case, the description of the system requires as many independent measurements as the number of levels. To make an example, the knowledge of the firms' specialisation in Prato does not in itself reveal the capabilities or the structure of the value chain of the system Prato, as these are a by-product of the reconfiguration mechanism and project-based dynamic introduced in section 3.3.1.4 (and in the introduction in section 1.6),

552 For instance see the classifications used in this research, i.e. the Cannari & Signorini and the Sforzi/Istat described in sections 4.1.3.1 and 4.1.3.2

which are inter-level coupling mechanisms, not knowable by firm's level analysis. This may imply that complex systems may require different or additional methods of analysis based on different methodologies than more linear systems. What is valid for A2, A1, and D2 may not be valid for D1. A measure of distribution such as the scale-free or self-organised criticality bridges the two levels and sheds some light on these mechanisms.

Second, the marked difference between A1, A2, and D2 on the one hand and D1 on the other hand suggests that the self-organisational dynamic of industrial clusters shows a marked acceleration after a threshold. The clear difference between the two classes of results suggests that something similar to a phase transition is taking place between D2 and D1. This implies that although the agglomerations belonging to D2 shows some features of clusters in terms of the parameters described in section 4.1.3.2 and although they are usually classified as clusters in the current literature, the SLLs belonging to D2 are not clusters, as they seem to remain below the threshold of the phase transition suggested above. As the power law analysis indicates a positive correlation between cluster dynamic and self-organisation, it is natural to describe the phase transition suggested above as related to self-organisation. In conclusion the gap between the two classes of results suggests (although it does not demonstrate) that the social and economic life of type D1 agglomerations is structured around a different attractor from the one dominating the other categories and that a discontinuity separates the two attractors' basins of influence.

6.1.4 CONCLUSIONS

The evidence and discussion presented in this and previous chapter indicate that the "cluster effect" exists and can be measured. It also suggests that clusters are self-organised critical systems. This implies that:

1) The cluster behaves as a system and not as an aggregate and therefore that the study of clusters requires the use of non-reductionist approaches.

2) The cluster is a system in the dynamic state known as on the *edge of chaos*. Structures of order emerge at a systemic level and evolve in a co-evolutionary dance with external and internal structures and constraints. Self-organisation leads the process.

3) This study has further refined the taxonomy proposed by Cannari and Signorini, generally confirming the validity of their approach. However, the large distance that separates industrial clusters of type 1 from type 2 together with the closeness of type 2 clusters to the other two taxonomic groups suggests that 'real clusters' are only those of type D1.

Finally I would like to stress that, to my knowledge, this type of research has never been attempted before.

6.2 *DISCUSSION OF DIVERSITY RESULTS*

The discussion of results will follow the same order of the presentation of results in chapter 5.

Variety

The results in Figure 41 and Figure 42 indicate that variety seems to be inversely proportional to the cluster content of the category. Variety gives a measure of the number of different species present on average in each category of industrial agglomerations as measured by SIC codes. If we suppose that each species inhabits a particular industrial niche, then variety gives an idea of the amount of niches present in a territory. However, variety provides no indication with regard to: first, niches similarity, for instance whether niches are all in the textile sector or spread across multiple sectors; and second, niches interconnection. Industrial agglomerations of type D1 show a double phenomenon of concentration: first, there are fewer niches than non-cluster and second, each niche is populated by a higher number of firms[553].

Distance

Figure 43 provides a measure of disparity, that is, the difference between species. Clusters (D1) show a distance considerably smaller than non-clusters, or in other terms, cluster's activities have a higher degree of similarity. This is related to the concentration effect already revealed by variety.

Generic diversity

---

[553] this is automatically true in the case of equal samples of 600 entries per SLL. Fewer niches (SIC) imply that the entries will be spread on a smaller spectrum of niches. However, the same result emerges

The measure of diversity in Figure 44 shows that diversity is inversely proportional to the cluster form. This seems to run counter to the theory elaborated in chapter 3 on the economies of diversity, which hypothesised a direct relationship between amount of diversity and self-organisational dynamic. As clusters were identified with the embodiment of self-organising social systems, the lack of evidence about the correlation diversity vs. self-organisation indicates that either the hypothesis is wrong or diversity requires a more complex analysis. However, the definition of diversity assumed in this study, although providing a sophisticated approach to the problem of measuring diversity, does not explore the issue of the interconnection among the elements composing the systems. From this point of view, the facts that aggregates present a more diversified spread of activities is not surprising. On the contrary, the lack of interdependence among elements can be seen as a lack of top-down constraints on the elements of the aggregates. Consequently as each element develops unconstrained from the others, overall diversity is higher.

A recurrent theme of the previous three figures is the polarisation of the measures between D1 and the other three categories. This point is discussed in section 6.1.3. The fact that D1 stand isolated suggests that a different dynamic is in action. This is the object of the analysis of systemic diversity.

Systemic diversity

The results of variety, diversity and systemic diversity are reported all together in arbitrary units in Figure 50. The comparison reveals contrasting trends.

- Variety and diversity correlate inversely with cluster content
- Systemic diversity instead correlate positively with cluster content
- D1 stands on its own when compared to the other categories

---

even when the total number of entries (without filtering the population of SLLs down to the sample of 600 entries) is taken into consideration.

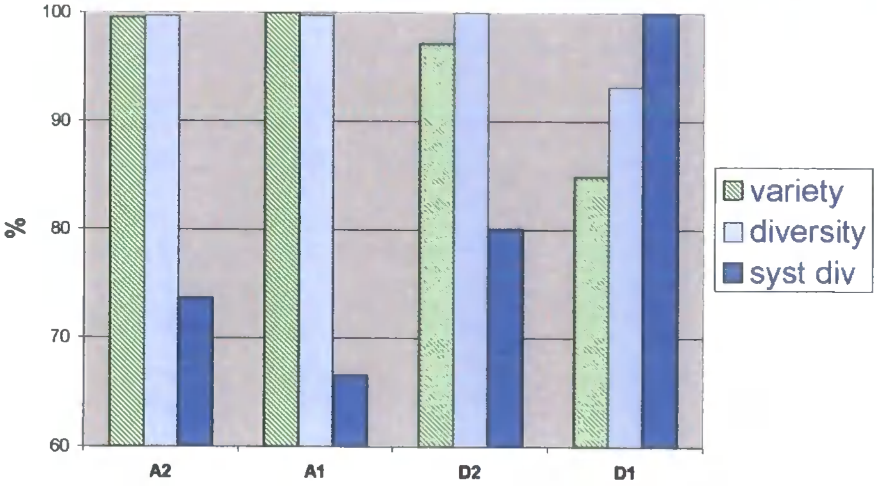


Figure 50: Variety, diversity and systemic diversity

What is the meaning of these contrasting trends?

First, they confirm that looking at diversity in isolation from the connectivity elements risks to give a false picture. High generic diversity is associated with aggregates of parts but high systemic diversity is instead associated with inter-connected and self-organising systems.

From a computational point of view the fact that higher systemic diversity is associated with systems characterised by lower variety and diversity is possible because the connectivity weight acts as a multiplicative factor that modulates the contribution of pairs of firms to systemic diversity. The diversity contribution coming from pairs belonging to species, which (according to the national average) don't interact financially, are totally excluded. Only pairs of interactive organisations enter with their appropriate weight in the computation of systemic diversity.

From a more general point of view, the systemic diversity results indicate that the proposition regarding the product of diversity and connectivity is confirmed. Self-organising industrial agglomerations are structures that show a higher product of internal diversity and connectivity when compared to the other types of industrial agglomerations.

However, although the proposition is reasonably supported by the available evidence, the latter is still too limited to definitely demonstrate the proposition. In fact, apart from the limitations discussed in the methodology chapter (section 4.3.2), the sample on which the analysis is based (ten SLLs per category) is relatively small. Moreover, as the selection criteria of SLLs are based on social and geographic considerations related to self-contained home-to-work commuting and as these are independent from the definition criteria of industrial clusters, the conclusion follows that the unit of analysis, which this research has adopted - the SLL – is sub-optimal. This term indicates that the SLLs are not designed having in mind the idea of carrying out research on industrial clusters or (at least directly) even on industrial agglomerations *per se*. As a result of this, the SLLs as units of analysis constitute a 'noisy' sample. The noise is generated essentially (but not only) by the lack of overlap between the boundaries of industrial clusters and SLLs and by the inclusion of every type of firms in the database[554]. If on the one hand this turns out to be beneficial, as it provides an unbiased classification sample for the analysis and some clearly defined boundary conditions, on the other hand, the high level of noise necessitates a really strong signal to emerge from the background and makes the exact evaluation of the differences between categories more difficult to estimate. For instance, an effect of this noise is likely to be seen in the high values of the standard deviations (see Table 12) associated with the measurements of diversity.

On the other hand, all the results (from power law, variety, diversity to generic diversity) converge to the same trend, that is the positive correlation between self-organisation, power law distribution and maximisation of systemic diversity. The general agreement in the trends among different measurements indicates that the argument is substantially correct. This conclusion is further confirmed by the fact that most of the systematic errors by which the research is affected tend to underestimate the internal diversity of clusters more than the other categories. For instance, the use of input-output tables, based on national averages, is certainly more appropriate in the case of industrial agglomerations closer to the national average (A2, A1 and D2), than for the case of the highly idiosyncratic industrial clusters. The fact that the results shown above for A2, A1 and D2 are invariably closer to one another than they are for D1 gives additional weight to the argument.

	Average				Std dev			
	A2	A1	D2	D1	A2	A1	D2	D1
Variety	1.12E+02	1.13E+02	1.10E+02	9.58E+01	1.08E+01	1.22E+01	6.57E+00	1.33E+01
Diversity	6.31E+05	6.32E+05	6.33E+05	5.90E+05	9.77E+03	9.99E+03	1.12E+04	5.79E+04
Syst div	2.47E+07	2.23E+07	2.69E+07	3.36E+07	9.39E+06	3.76E+06	6.26E+06	1.92E+07

Table 12: Averages and standard deviations of variety, diversity and systemic diversity [arbitrary units]

The validation of the proposition on maximisation of the product diversity-connectivity changes the way in which we see industrial clusters. The traditional view considers clusters as a geographic concentration of companies specialised in the same sector in relationships of competition/cooperation among them. According to this view, the reason behind the success of clusters lies with a series of factors, namely: the cooperative element, the low costs associated with low tech productions, the incredible commitment and motivation associated with the enlarged family ownership in the context of a cohesive community, and finally, the chameleontic capacity of adaptation typical of small businesses. In short, a cluster is not qualitatively different from a network of small firms; with the only difference that the geographic concentration of firms magnifies the characteristics of competition and cooperation already present in any network.

This 'reductionist' view is in conflict with the picture that emerges from my research and from the literature that describe clusters as complex systems. According to this latter view, the actions of exploration and exploitation that the agents carry out in parallel every day produce a very elaborated order at least as complex as any central planner or collection of parts could ever produce. The spontaneous order established among autonomous actions is due to the action of multiple feedback channels. These constrain the actions of the agents toward acts of exploration and exploitation that are resonant (and path-dependent) with the actions of surrounding agents. The nature of the feedback channels and the type of constraints and possibilities it creates for agents have been introduced in chapter 3 (section 3.3.1 and subsections) under the name of economies of diversity. The increasing returns nature associated with economies of diversity coupled with the dissipative nature of clusters makes the survival of cluster depending upon a continuous process of change and exploration in which the structure adapts to the changes brought forward by endogenous and environmental transformation processes. The loops of the economies of diversity expand the sphere of exploration along specific trajectories influenced by the constitutive processes of the loops. Each loop adds new features to existing diversity as represented in the figure[555] below.

---

[554] More on this point in section 4.1.3.3

[555] the figure is introduced in section 3.3.2

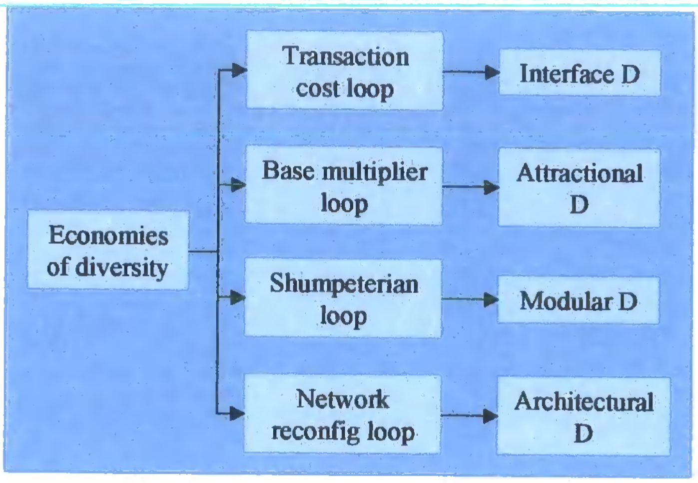


Figure 51: The contribution to diversity of the economies of diversity loops

The validation of the proposition suggests that the expansion of diversity takes place within the context of connectivity of the cluster. The proposition singles out clusters as systems where the process of exploration of the solution space is strongly affected by the structure and nature of links among agents. The validation of the two propositions on power law and maximisation of diversity-connectivity constitutes a strong indication that the dynamic of self-organising systems is radically different from the evolution of hetero-organised systems on the one hand and from the dynamic of aggregates of parts. The cluster form bases its survival on distributed intelligence, emergent reactions and self-organisation, all terms captured in the economies of diversity framework. The natural form of organisation is the decentralised network. Hetero-organised systems instead survive by centralising intelligence, implementing control by a set of rules and reducing diversity. In the case of aggregates, the absence of internal constraints does not give rise to either dynamic. To sum up: in the first case, self-organisation leads to a network structure, power law distribution, and maximisation of diversity-connectivity; in the second case, hetero-organisation leads to centralisation, process efficiency, minimisation of diversity and streamlined connectivity; in the third case, random dynamic leads to decentralised distribution of parts, lack of structure, maximisation of generic diversity, low systemic diversity and finally to a bell (in its pure form) distribution of relevant variables.

However, the question remains open whether the transition between self-organised and hetero-organised systems is continuous or discontinuous. A different type of discontinuity may exist at the interface between aggregates and systems. As far as this research is concerned, the fact that the industrial cluster D1 category stands on every single measure in isolation from the others seems to lend support to the idea that D1 is subject to a specific dynamic probably triggered by some sort of bifurcation (or symmetry breaking) and phase transition. Once started, the self-sustaining dynamic of the economies of diversity reinforces the dominant logic of distributed self-organising systems. This observation raises a difficult issue: if it is true that diversity and connectivity starts a self-reinforcing dynamic after a threshold, can that threshold be determined and how? I'll suggest in the chapter on further research some suggestions to tackle this issue.

6.2.1 THE PRATO CASE

Finally, Figure 52 and Table 10 show the main result of variety, diversity and systemic diversity in the case of the most famous industrial cluster, the textile cluster of Prato. The results give further support to the proposition, even though the contrast this time is not between different geographic agglomerations, but instead between clustered and non-clustered activities within the same cluster. The contrast is stark. Systemic diversity relative to core cluster activities is five times larger than non-clusterised activities. The difference indicates that the contrast between the dynamic of self-organising networks and aggregates can coexist side to side within the same territory.

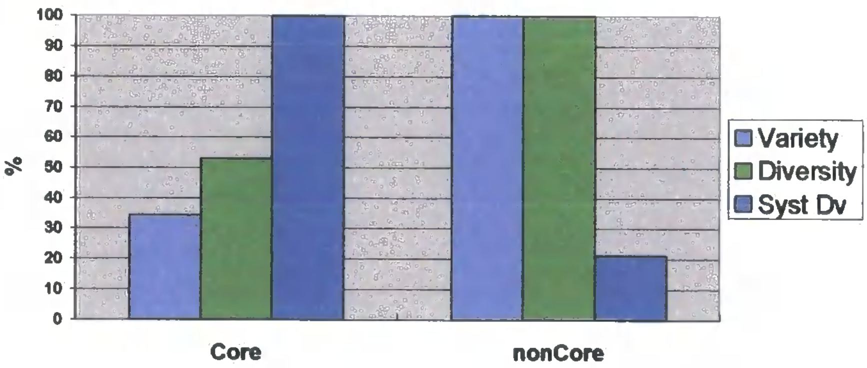


Figure 52: Variety, diversity and systemic diversity between core and non-core cluster activities in Prato

What are the final results of this research? Clusters constitute a specific type of econosphere, whose driving principles are self-organisation, economies of diversity and a configuration that optimises the exploration of diversity starting from the configuration of connectivity of the system. The maximisation of diversity and connectivity takes place "*without destroying the accumulated propagating organisation that is the basis and* *nexus from which further novelty is discovered and incorporated into the propagating organisation*"[556]. The structure of the propagating organisation obeys a power law distribution. This distribution maximises the exploration of nodal features, connectivity patterns and system wide behaviours.

These results confirm the general intuition by Kauffman that the evolution of biosphere and econosphere is subjected to some general 'laws'[557] that cut through the specifics of individual systems. Power laws and the idea that organisations are propagating set of autocatalytic processes expanding diversity along trajectories set by environmental pressures and internal connectivity are two of these 'laws'. My research develops a specific formulation for these two laws in the economic context of industrial agglomerations and demonstrates their validity.

---

[556] (Kauffman 2000) p.85

[557] I am not talking of laws with a capital L, such as those prevailing in physics and chemistry, but more about dynamic patterns which may be subject to mathematical formalisation

Chapter 7: Conclusions

7.1 THE THESIS IN BRIEF

Complexity theory is a new interpretative framework that has the potential of leaving a mark on the evolution of thinking in natural and social sciences. Complexity is an 'organismic' framework. Its underlying philosophy (its *weltanschaung)* rests on different assumptions from the dominant neoclassical theories in natural sciences and economics. It tries to explain the origin of order, the robustness and the decay (or catastrophic demise) of social and natural systems thanks to a few universal laws, which are more similar to dynamic trajectories of development than to the prescriptive laws of Newtonian sciences. Complexity is also a multilevel theory. Natural and social phenomena are the result of the aggregation of agents at multiple levels. Individuals form families and cliques whose behaviour is directly regulated by culturally acquired traits and specific self-reinforcing interactions valid only for those groups. Families and groups form villages and cities. At this level, the rules or interpretative frameworks valid for individuals are of limited validity in the interpretation of social and economic phenomena at the level of cities. For instance the emergence of homogenous neighbourhoods is very often unrelated to individuals' intentional behaviour. The result of those interactions though is the formation of macro-patterns, which we call edge cities. Cities acquire a specific character, and although they consist of different elements - individuals, groups, neighbourhoods, physical artefacts, buildings, roads and history - none of these aspects can explain their character. Again as the formation of a neighbourhood is often not driven by intentional behaviour (for instance racial discrimination), so the character of a city is only loosely matched to the character of its inhabitants and institutions. The aggregation of higher-level structures (with the consequent emergence of new systems) continues upward toward regions and nation states. At each level of aggregations new laws and new fields of endeavour emerge.

Chapter 1: The Story of Meadow

I've tried in the first chapter to tell *in anger* the story of the formation of one such system. Meadow, an isolated village based on "*backyard capitalism*" which evolves into a sophisticated cluster characterised by a complex self-regulatory production and innovation system, is the story of the passage of a system through a series of phase transitions (or bifurcations) each of which gives rise to the emergence of a different system structured around different dynamics. In short, it is the story of the emergence of order in a social system. The creation of order does not follow a predetermined trajectory set by universal laws, as in classical physics and neoclassical economics, in which the attainment of equilibrium wipes out the system's idiosyncratic aspects that emerged in the pre-equilibrium phase. The Laplacian dream of a perfectly computable universe, and therefore predictable by anyone who could analyse it using the language in which the universe is written, disappears. From the point of view of predictability, Meadow and many industrial clusters seem to behave like systems characterised by maximum algorithmic complexity[558]. However to renounce a Laplacian world does not imply the end of theorising but only its recalibration.

Chapter 2: Literature review

The example of Meadow is interesting as it posits the question: "If systems based on self-organisation emerge via semi-spontaneous creation of order in an endogenous way, then what theory do we need to make sense of emergence?" The phenomenon of emergence is difficult to analyse via theories that assume that the behaviour of the system can be reduced to a sum of its components' behaviour. In Meadow the emergence of project-based innovation or the mutual closure of autocatalytic loops generated a business and social environment, whereby bottom-up dynamics coexisted with top-down constraints. The interaction of both (bottom-up and top-down) created a new organisational level whose rules were not amenable to the ones of agents at the level below. The emergence of new organisational levels (with the consequent emergence of new dynamic attractors) creates a discontinuity in the system's history between the agents and the higher level created by the aggregation of the agents. Although the new organisational levels are formed by lower level components and are therefore coupled to the previous levels, their dynamics are partially decoupled[559]. Reductionism can work horizontally along levels but not diagonally across levels. The theory we need is one that a) takes account of the inherent limitation that emergence imposes on the

[558] Software programmes that show no pattern of regularities are systems with maximum algorithmic complexity[558]. The absence of any recognisable pattern among the software lines makes it impossible to devise a shorter code able to capture the future behaviour of the system (Gell-Mann 1994). The only way to understand how the software runs is to let it run. Likewise the only way to understand an industrial cluster is to observe it evolving without expecting to find ways to generally describe the evolution of the system under any possible circumstances.

[559] (Nicolis and Prigogine 1989)

knowledgeability of systems and b) embeds the emergence of new, more complex organisational levels into a more general theory that includes reductionism as a particular case.

I have introduced some of the principles of complexity theory and very briefly sketched their historic evolution. Complexity as a theory of emergence in self-organising networks is still young and the fields to which it has successfully been applied are limited. In the case of industrial clusters there are very few articles directly inspired by complexity. Only recently there has been an increase in interest. A recent conference[560] organised in Italy, called "*Complexity and Industrial Clusters*", has probably been the first conference to gather economists, sociologists, economic geographers and physicists, in order to discuss explicitly the application and relevance of complexity to industrial clusters. This is somewhat surprising. Industrial clusters are an ideal field of application for complexity inspired theories and models. As seen in the case of Meadow, clusters are self-organising networks with distributed control and emergent behaviours. A thorough study of the traditional literature on clusters, the authors often cited in this work including the "Becattinis", the "Storpers", the "Saxenians", etc., reveals that these authors find traditional approaches to the study of economic and social agglomerations insufficient, and indeed, many of them have invented new approaches (see section 2.2 and subsections), which under many points of view reflect an implicit acceptance of complexity principles. These frameworks make frequent use of explanations that exhibit a self-sustaining and often circular logic, they describe situations in which locality of agents' actions yield non-local macro effects and so on. Emergence though and the principles of self-organisation are not explicitly recognised. My research moves into the direction of filling some of these voids. The first objective has then been to apply complexity theory to the world of industrial clusters.

Chapter 3: Theoretical frameworks

My research started from a simple observation: industrial clusters are places where a set of autonomous organisations, in absence of external coordination mechanisms, achieves a high degree of coherent action. A few questions come to mind. First, what is the origin of order that an industrial cluster exhibits? Second, from a more operational

---

[560] "Complexity and Industrial Clusters: Dynamics and Models in Theory and Practice". A Conference Organized by Fondazione Montedison Under the Aegis of Accademia Nazionale dei Lincei, Milan, June 19-20th, 2001

point of view, how can one recognise systems characterised by emergent order from other types of systems? Third, what are the distinctive features that set self-organising industrial clusters apart from other types of industrial agglomerations?

The first question is dealt in the literature review and in a more complete way in the theoretical chapter. I have shown in the imaginary story of Meadow how a system could structure itself around some emerging dynamic processes and sets of environmental opportunities. The process is essentially endogenous in the sense that the generation of inter and intra organisational routines emerges in the dialogue between on the one hand the local conditions of connectivity, behavioural rules and complementarities among agents, and on the other hand, the need to adapt to changing environmental conditions and/or exploit opportunities. This generates systems that are informationally open but organisationally closed. Although the trajectory of evolution of the system may tend to general equilibrium conditions in presence of long-term macro and micro environmental stability, this does not happen when waves of innovations follow one after the other. It has long been suspected that the evolution of systems does not proceed with Darwinian incrementalism, on the contrary it alternates periods of environmental stability with short bursts of Cuvierian radical changes:

"*...these repeated [advances] and retreats of the sea have neither been slow nor gradual: most of the catastrophes which have occasioned them have been sudden ...*"[561]

Complexity has equipped us with theoretical instruments to understand the transition from one period to the next. We have understood (or at least there is mounting evidence) that the transition needs not necessarily be driven by a catastrophic event (the Yucatan asteroids for the end of the dinosaurs or the Napoleonic army for the ascent of the liberal state) but can result from the endogenous accumulation of micro-fractures that result in sudden large-scale events. We also know[562] that the transition really corresponds to the passage from a set of dynamic attractors to another. These evolutionary systems do self-organise around the attractors but the results of self-organisation depend on the specific attractors. We also know that the emergence of new attractors (for instance the transition from the hunters-gatherers to the agricultural society) happens through one or multiple bifurcation in which random circumstances play an important role.

---

561 (Raup 1999) p.30

562 (Allen 2001)

In the theory chapter I have presented a sketch of a theory that describes the crystallisation of industrial clusters around some general dynamic patterns that act as attractors. The fact that the evolutionary dynamic of systems building structures of order around new attractors takes place rapidly (at least compared with the period of incremental change that ensues) implies that the elaboration of the new system must be based on mutually self-reinforcing growth and selection mechanisms. If this were not true, incrementalism would lead to a smooth landscape of undifferentiated hills, whereby each hill would be similar to its neighbours and differences would gradually increase with distance. This is not what we observe. The impressive agglomerations and concentrations of sectors in which turbulent change is the norm indicate a world dominated by increasing returns. I have in the chapter on theory presented an architecture of different mechanisms based on increasing returns. The mechanisms are organised as autocatalytic loops that tend to recreate and reinforce themselves.

The initial point concerns the relationship between diversity and growth "… *diversity probably begets diversity; hence diversity may help beget growth*"[563]. Diversity is autocatalytic, it leads to further diversity. The examples, given in the literature review (section 2.5.2) showing the self-reinforcing dynamic in the development of agriculture from the hunters-gatherers society, constitute valid evidence of the mechanism. The expansion of diversity is likely to result in the closure of a catalytic cycle. The closure causes a series of effects: a) it determines the emergence of an internal architecture of connectivity based on dynamic self-sustaining processes, b) it provides a preferential path of development by providing rules for selection and exclusion of processes, c) it introduces an internal dimension and dynamic, thereby giving rise to the emergence of a unique identity. In short, closure introduces internal rules of organisation, which causes the transition from an aggregate of parts into a system and the emergence of a coevolutive dynamic. Coevolution and closure are conjugated variables, which act to decouple the system from its environment. The system is open to external information and energy, which are used to construct complex internal structures, whose organisational principles are however internally determined. The rules of cognition and engagement with the environment are fixed by the internal dynamic. In other terms the organisational rules are genotypic, whereas the space of implementation of the rules is phenotypic.

Eigen and Schuster[564] showed that the closure of catalytic reaction formed the origin of systems. Kauffman showed that diversity is critical in achieving closure. The rules that generate order in the transition from aggregates to industrial clusters are the mutualistic autocatalytic loops that I have presented in the theory chapter. Each loop finds its roots in the interaction of diverse parts and stimulates the generation of further diversity. Systemic diversity is different from diversity prevailing in aggregates. In fact the former is the result of the expansion of the system in the *adjacent possible*, constrained by path-dependency and occurring along a number of technological trajectories, the latter is merely fortuitous. In other words the expansion of diversity in systems is a result of entangled processes. The entanglement is an effect of the connectivity of the system and expresses the idea that the actions of the agents can not be separated from the other agents' actions and from the environment. Briefly, in self-organising systems, diversity and connectivity are inextricably linked, whereas this is not the case with aggregates. The identity and evolution of industrial clusters depend on a set of interlinked autocatalytic loops whose basic configuration emerges at some critical bifurcation moments. This provides a tentative answer to the first general question regarding the origin of order in industrial clusters.

The second question – how to distinguish industrial clusters from other types of agglomerations – leads into the search for structural characteristics that can be used as indicators of the specific dynamics described above. The starting point is that industrial clusters as self-organising systems enjoy a set of properties: they consist of networks of autonomous agents; their pattern of connectivity exhibit weak and strong linkages; and they obey power law distributions. The three features are connected. In fact, in self-organising systems chaotic and ordered features coexist. Chaotic dynamic allows for frequent reconfiguration and emergence of novelties, ordered dynamic allows for robustness and homeostasis. The balance between the two is described by the metaphorical expression of *edge of chaos*. At the *edge of chaos* the distribution of connectivity alternates strong links within highly connected sub-networks with weak links connecting the various local networks (within the system). The distribution of links within and across sub-networks and the dimensions of the highly connected sub-networks obey a power law. The emergence of a power law is a sign of a self-organising

---

[563] (Kauffman 1995) p.292

[564] see (Kauffman 1995)

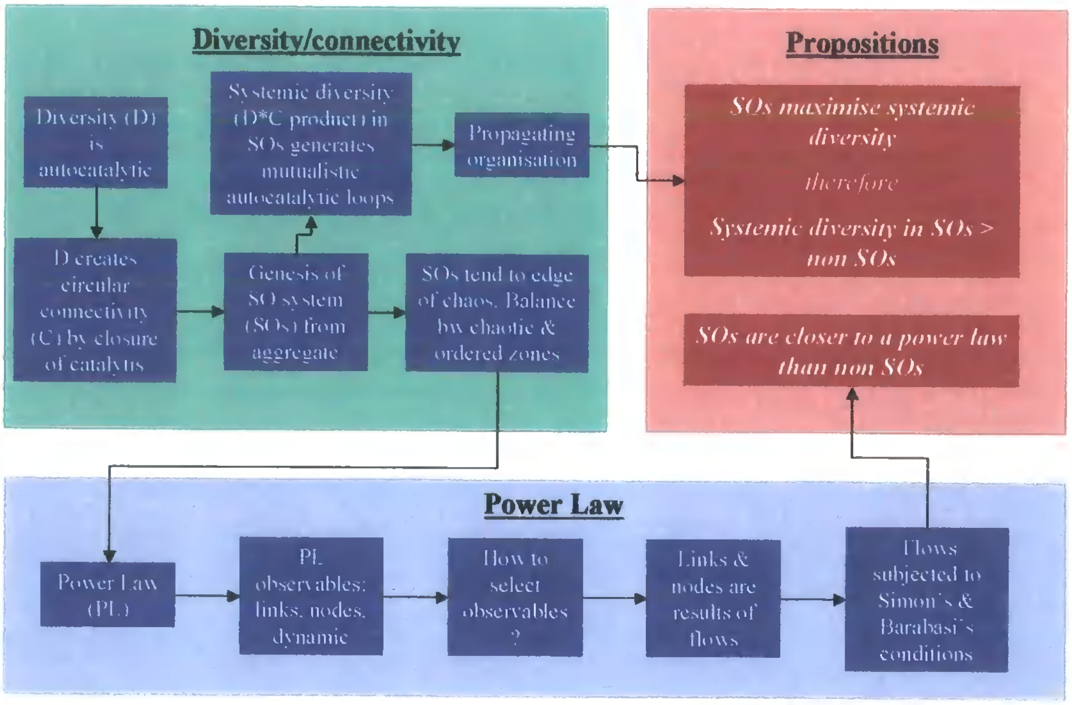
dynamic. In fact the power law indicates that the system tunes itself toward a state whereas the distribution of structural and behavioural properties is such that all scales and behaviours of relevant network's properties are present in the system and follow a simple mathematical distribution

Power laws emerge in the three constituent elements of networks, i.e. links, nodes and behaviour. I have presented the argument that the occurrence of power law for these three elements is the manifestation of the same underlying dynamic, that is, the self-organising nature of networks at the *edge of chaos*. I have also suggested that a transformation of variables changes one variable into the other, thereby raising the possibility of using any variable for the determination of the structural and/or dynamic state of the system. In short, self-organisation is explainable in terms of the set of interdependencies that determine the set of flows between nodes and links. This is something that a statistical analysis on appropriate systems can test. If self-organisation and power law are correlated then we can use power laws to reveal self-organisation in action. From this point of view, a comparison between systems characterised by different degree of self-organisation should reveal a correlation between self-organisation and closeness to power law.

The theory summarised above was used to generate two propositions:

1. Self-organising systems are closer to a power law than hierarchical systems or aggregates (collection of parts). For industrial agglomerations (SLLs), the closeness to a power law is related to the degree of self-organisation present in the agglomeration, and emerges in the agglomeration's structural and/or behavioural properties subject to self-organising dynamic.
2. Self-organising systems maximise the product of diversity times connectivity at a rate higher than hierarchical systems

The logical steps regarding the translation of the theory into propositions are represented below.



Chapter 4: Methodology

In the chapter on methodology I have first, discussed the methods of research adopted to test the two main propositions, provided a rationale for those methods and illustrated the limits of the research methods; second, I have introduced the units of analysis adopted in my research and provided a rationale for their selection; and third, I have exposed in necessary detail the procedures and steps taken to validate the propositions. This chapter offers two main contributions. First, I propose and implement a method to measure the effect of self-organisation in industrial agglomerations by using a power law distribution. Second, I propose and develop a method to measure generic and systemic diversity in industrial agglomeration.

Chapter 5 and 6: Results and discussion

The empirical work, carried out on Census data relative to different types industrial agglomerations in Italy, confirms the validity of the propositions. The picture that emerges is as follows:

3. A power law analysis is an effective instrument to discriminate between systems driven by self-organising dynamics and systems driven by hetero-organised dynamics (aggregates of parts and hierarchical systems). In the context of the Italian industrial agglomerations my results confirm that the self-organising dynamics that characterise the presence of a cluster can be revealed by a power law analysis. Therefore my

suggestion that a power law is an indicator of the emergence of a system from an aggregate of parts is confirmed by the empirical data. Moreover, as industrial agglomeration are a complex mix of self and hetero-organised dynamics, I show that a power law analysis is able by making use of a regression analysis and, at least in the first approximation, to indicate the weight of the different dynamics.

4. The point made above together with first, the idea (presented in section 3.2.1) regarding the equivalence of the different power law theories and second, the discussion of the correlation between self-organised criticality and edge of chaos dynamic (section 3.2.3), suggests that power law distribution is an indicator of an *edge of chaos* dynamic. Industrial clusters then are social systems on the *edge of chaos*. When such a dynamic arises, the system shows a power law distribution of chaotic and ordered features. As this type of distribution doesn't have a typical scale, this means that industrial clusters thrive and survive by exploring a very broad range of possible dynamics.

5. The previous point leads me to the second part of the results. The origin of the self-organising dynamic, revealed by the power law analysis, lies with the capability of exploring diversity. My results confirm that self-organising systems maximise a particular type of diversity, that I called systemic diversity. I propose to measure systemic diversity by the product of diversity (as measured by a mix of three parameters, variety, balance and disparity) and internal connectivity. I also propose to use the concept of systemic diversity as a way to operationalise Kauffman's concept of the *propagating organisation*[565].

6. Specifically, my research demonstrates that industrial agglomerations closer to the cluster form shows a higher systemic diversity than systems more dominated by aggregations or hierarchical organisational forms. Moreover, in terms of the characteristics of the Italian industrial agglomerations this study has further refined the taxonomy proposed by Cannari and Signorini[566] and generally confirms the validity of their approach. However, the large distance that separates industrial agglomerations of type D1 (so called super-clusters) from agglomerations of type D2 (usually classified as clusters), A2 and A1 suggests that 'real clusters' are only those

---

565 see next section

566 (Cannari and Signorini 2000)

of type D1. This result seems also in line with the autocatalytic nature of cluster formation and development, which would exclude a continuous distribution of agglomeration types between the cluster and non-cluster.

What are the final results of this research? Clusters constitute a specific type of econosphere, whose driving principles are self-organisation, economies of diversity and a configuration that optimises the exploration of diversity starting from the configuration of connectivity of the system. The maximisation of diversity and connectivity takes place "*without destroying the accumulated propagating organisation that is the basis and nexus from which further novelty is discovered and incorporated into the propagating organisation*"[567]. The structure of the propagating organisation obeys a power law distribution. This distribution maximises the exploration of nodal features, connectivity patterns and system wide behaviours.

These results confirm the general intuition by Kauffman that the evolution of biosphere and econosphere is subjected to some general 'laws'[568] that cut through the specifics of individual systems. Power laws and the idea that organisations are propagating set of autocatalytic processes expanding diversity along trajectories set by environmental pressures and internal connectivity are two of these 'laws'. My research develops a specific formulation for these two laws in the economic context of industrial agglomerations and demonstrates their validity.

7.2 IMPLICATIONS

The theory on industrial clusters presented in chapter 3 and the demonstration of the validity of the two general propositions related to that theory have a wider significance that extends beyond the specific case of industrial agglomerations in Italy. I will develop two implications of this. First, I present a critique of the economic geography discussion regarding localisation economies vs. urbanisation economies. My critique is based on the central argument of my thesis that diversity and connectivity are conjugated variables. In the case of the search for the enabling conditions of innovation, this approach implies that the focus on diversity alone runs the risk of neglecting the complexity of the problem

---

[567] (Kauffman 2000) p.85

[568] I am not talking of laws with a capital L, such as those prevailing in physics and chemistry, but more about dynamic patterns which may be subject to mathematical formalisation

and consequently to develop the wrong correlations. I present an example of such an approach. Second, I develop some further properties of the *propagating organisations* by showing that these systems have evolutionary properties that make the evolvability of such systems speedier.

7.2.1 DIVERSITY AND INNOVATION: LOCALISATION AND URBANISATION ECONOMIES

The distinction between localisation economies and urbanisation economies has been a classical terrain of academic dispute since the 60s, when Jacobs[569] in her seminal work on cities advanced the hypothesis that cities are cradles of innovation as they have a higher interval economic and social variety than villages. The thesis did not attract much attention until the 80s, when the rediscovery of industrial clusters and of the amazing concentration of innovation spots in cities forced the issue of the sources and enablers of innovation.

Localisation economies refer to the pattern of externalities associated with the geographic concentration of firms in the same or very closely related sectors. Localisation leads to the simultaneous presence of multiple firms specialising in a limited spectrum of activities. Proximity and similarity of knowledge and production techniques speed up the diffusion and adoption of technologies, once the usefulness of those has been demonstrated within the locality. Hence localisation economies cause increasing urban specialisation[570] and consistent economies of scale resulting from the sharing of labour markets, infrastructures, and equipment[571].

Urbanisation economies instead "*reflect externalities associated with the simultaneous presence of firms from various industries, extensive infrastructure or a large pool of labour in a given location*"[572]. The co-location of different firms stimulates innovation through the exchange of ideas and borrowing of processes/knowledge from related but different sectors. As Marshall[573] noticed: "*inventions and improvements in machinery ... have their merits promptly discussed; if one man starts a new idea, it is*

[569] (Jacobs 1967)

[570] (Sabel 2001)

[571] (Harrison and Kelley 1996) p.7

[572] (Harrison and Kelley 1996) p.6

[573] cited in (Kelley and Helper 1997) p.12

taken up by others and combined with suggestions of their own; and thus it becomes the source of new ideas". Although the Marshallian mechanism works also for localisation economies, it becomes orders of magnitude more effective in presence of urbanisation economies.

A branch of economic geography[574] has dedicated a substantial effort to the resolution of the question whether specialisation is more important than diversity in promoting innovation. The stunning observation by Feldman and Audretch[575] (studying a sample of almost 4000 US firms in 1982) that 96% of innovations happened to be localised in metropolitan areas comprising less than 30% of the overall population, constitutes a strong evidence in favour of the 'Jacobs' mechanism[576]. The studies cited surveyed large statistical data regarding populations of firms or agglomerations at the county level using indicators of the type used in this research (although the expression used to measure diversity, the Hirshman-Herfindhal[577] seriously distort the measurement of diversity as it takes into account only variety and balance, but neglect disparity) in order to ascertain the role that specialisation (localisation economies) and diversity (urbanisation economies) play as enablers of innovation.

Some authors arrive to the following conclusions:

"*Where characteristics of locale hypothesized to be sources of potential agglomeration economies are significant, it is diversity, reflecting the properties of urbanization, more consistently than sameness of location, that appears to be the motivating factor*"[578]

They also add:

"*Thus we add our voices to those who have found that the 'Jacobs' effects – in a word diversity – are likely to be sources of dynamic external economies at least as important as the presence of many other firms or workers in the same industry or sectors; the variable emphasised in the industrial district literature. Urbanization, or diversity, is as important as, or more important than, localization, or sameness, in contribution to an explanation of which kind of firms adopt new technologies and which are less likely to do so*"[579]

---

[574] (Harrison and Kelley 1996; Kelley and Helper 1997; Duranton and Puga 1999; Duranton and Puga 2001; Sabel 2001)

[575] (Duranton and Puga 2001) p.8

[576] Further evidence come from (Harrison and Kelley 1996; Kelley and Helper 1997; Duranton and Puga 1999; Duranton and Puga 2001; Sabel 2001)

[577] (Duranton and Puga 1999) p.4

[578] (Harrison and Kelley 1996) p.24

[579] (Harrison and Kelley 1996) p.24

However why and how diversity promotes a more innovative environment remains a mystery. This branch of economic geography literature delves into the properties of diffusibility of tacit knowledge and co-location. It introduces distinctions based on life-cycle models, whereby diversity would play a more important role than specialisation when the necessity to "*prototype*[580] technical solutions to new problems in absence of pre-existing knowledge and "*compare*"[581] them across diverse sectors is dominant, that is at the beginning of the a new technological cycle (we find again the contrast between 'young'[582] firms which prosper in a diverse environment and 'mature' firms which instead profit from specialisation in the pursuit of economies of scale and efficiency). It stresses dynamic externalities in opposition of Krugmanian static economies of scale.

As Charles Sabel notices referring to the evidence cited above: "*by themselves these results are hard to interpret. They could be an artefact of aggregation. It is hard to imagine, in the absence of any account of the micro-mechanisms at work, just how the diverse airs of a great metropolis or suburban county foster innovation. It is equally hard to see why proximity to cooperatively competitive neighbours (talking incessantly, it might seem, about their ingenious plans to outdo you) hinders it*"[583]

From the point of view of my research the limitations of the localisation vs. urbanisation economies lay not so much in the definition adopted for diversity (the definition that they use, the Hirshman-Herfindel does not take into proper account disparity), but rather in the assumption that diversity taken on its own can be correlated with innovation. In opposition to the results mentioned above, my research found out that agglomerations verging on the side of aggregates show a higher diversity than agglomerations characterised by more systemic features. Although I haven't directly correlated diversity with innovation, and therefore, I can not claim that 'systemic' agglomerations innovate more than 'aggregated' agglomerations, I can however point to a large body of literature which demonstrates the highly innovative capability of industrial clusters. The examples of Silicon Valley, Prato, the City and Hollywood speak loudly. Diversity on its own does not explain innovation. The argument of the

---

580 (Sabel 2001) p.18

581 (Sabel 2001) p.18

582 (Sabel 2001)

583 (Sabel 2001) p.18

urbanisation economies, which states that in a diverse environment it is easier to 'borrow' knowledge from one sector and apply it to another, presupposes the existence of channels along which knowledge can be transferred. That is, it hides the assumption that some forms of linkages connect complementary techniques and knowledge residing in different entities. However, connectivity can not be taken for granted. Agglomerations where diverse communities are segregated and separated by invisible walls contribute to diversity not to connectivity. In large urban agglomerations fragmented into edge cities where little or no interaction takes place local communities, the Hirshman-Herfindel measure would find a net, probably very high, amount of diversity, but flows of knowledge and information would be remarkably small, if not altogether absent, thereby preventing the establishment of the fertile conditions that permit the conversion of diversity into innovations. Once the issue of the linkages between diverse organisations is taken into account, either of the 'weak' or 'strong' type, the discussion regarding localisation or urbanisation economies changes. In fact, if a measure of systemic diversity were carried out on an urban agglomeration dominated by segregated edge cities, the net result of systemic diversity would probably be rather low. This observation leads me to a further point. The research concerning the micro-economic causes that link a diverse environment to innovation addresses the question from a reductionist standpoint. These micro-economic causes interact with one another in a complex fashion that depends on the specific elements of each system. The interlocked loops that constitute the economies of diversity are formed by distinct elements, but these in isolation would not explain the positive feedback dynamic that characterise the concentration of innovation in a few distinct locations.

In synthesis, my critique of the localisation vs. urbanisation economies stems directly from the distinction I have introduced and applied in this thesis regarding the difference between generic diversity and systemic diversity. Seen from the point of view of this distinction (which is one of the contributions to knowledge of my thesis), I argue that the urbanisation vs. localisation economies discussion overlooks the point that innovation results not from diversity *per se*, which in itself is blind and inert, but instead from the expansion of diversity. Diversity becomes 'active' when it is organised into an architecture whose linkages are made by knowledge flows. Systemic diversity, that is, the product of diversity and connectivity is the important parameter.

7.2.2 DIVERSITY, CONNECTIVITY AND EVOLUTIONARY CONSIDERATIONS

The *propagating organisation* maximises systemic diversity. I have proposed in the chapter on theory a general set of dynamic loops that drive the expansion of the sphere of diversity. If the identity and structure of the *propagating organisation* is to be identified with an architecture of processes subject to evolutionary change, the relevant question arises as to what guarantees the robustness of the system. Biological entities defend their identity and dynamic structure by embedding them into an information code, the DNA. Conventional economic organisations do something similar by creating formal rules, codified procedures and legal structures.

The propagating organisation seems on paper to be less robust and potentially less able to survive than conventional organisations. However, at least in the case of industrial clusters, resilience to fluctuations in demand (or outright recessions) seems to be higher than in the case of large organisations[584]. The adaptability of clusters is remarkably high, but what is the origin of such capability? I think that the answer lies in the faster 'metabolism' that characterises the '*propagating organisation*'. Higher metabolism (or in other terms, faster coevolution[585]) leads to a better 'evolvability'[586], that is to a capability to develop robust evolutionary mechanisms that makes evolving easier and faster. Evolvability entails the capability to generate diversity, that is, to produce variations more likely to be beneficial for the whole system and at a rate higher than competitor forms. In the language of evolution this translates into faster adaptive radiation[587]. I argue that the self-organising properties of industrial clusters are at the root of evolvability.

First, self-organising systems can compartmentalise their internal structure. I have shown how this feature is tied for instance to a power law distribution. The great advantage of compartmentalisation is to isolate successful social and business mechanisms and to embed them into stable structures. The role of small organisations in a cluster is to shield the elements of competitive knowledge from rapid evolutionary

---

[584] (Becattini 1998) pp.128-131

[585] the terms is from McKelvey (McKelvey 2002)

[586] (Gerhard and Kirschner 1997)

[587] evolvability is connected to adaptive radiation, which refers to the capability to generate phenotypic change to occupy all available niches (see (Wilson 1992) and (Mayr 2001)).

change and to preserve them in the distributed organisational memory. The elements of the production techniques held at the level of individual organisations constitute the modules of the wider complex of operations that lead to the finished products. In the context of the organisational mechanisms introduced in chapter 3, these elements are the ones that result from the process of phase specialisation, according to which the emergence of divisibility in the way techniques are carried out takes place via a process of organisational partition. The newly formed firm and the old are now responsible for the indivisible elements of the techniques[588]. The compartmentalisation allows innovation to go on at the level of elementary production techniques, without directly affecting the recombination of the elementary techniques (modules) into rapidly changing production networks. In fact, the modules are owned by firms and play, using a biological metaphor, the roles that genes play in the transmission of the information. They are partially decoupled from the higher level, the network level, which instead emerges from the lower level interactions among firms. The rate of change of the two levels is very different. Elementary techniques, the 'genes', change (usually incrementally) at a much slower rate than recombinant networks[589]. The network, the higher-level agent that produces the final product, is modified each time a new network is formed around a new project.

The second important mechanism of the '*propagating organisation*' is redundancy. This is crucial to ensure that innovation does not stop at the level of modules. I have shown how redundancy leads to Shumpeterian competition. At the same time redundancy creates a multiplicity of channels along which parallel experimentation can take place. It also guarantees that experimentation with similar modules takes place simultaneously in different networks. Without redundancy, each network could only

---

[588] The idea that techniques play the genes role in evolution of technology is by Mokyr (Mokyr 1991). The dynamic of the process has been described in the Shumpeterian competition loop in chapter 3.

[589] the evolution of techniques shares many elements with the evolution of species and languages. In fact, species and languages evolve at a much faster rate than the evolution of the basic blocks composing biological species and languages. Most of the genes, as most of the letters composing our alphabet, have been around in their current form for millions (or thousand for languages) of years. Similarly fundamental techniques that constitute the building blocks of many technologies have hardly changed in the past hundred or so of years. Many basic techniques in mechanics, as for instance the universal joint, are still used. However the interaction of basic techniques in intricate architectures has given rise to complex systems. The distinction between basic techniques and complex technologies is reminiscent of the distinction between genes and species and the consequent discussion between genotypic and phenotypic evolution.

experiment with different modules, thereby reducing the space of possibilities at the architecture level.

Redundancy is strongly linked to the third mechanism, that is, weak linkages among lower level agents (firms). Weak linkages make possible that the emergence of regulation among agents in the formation of a network conserves a degree of freedom. An example of the contrary is the Japanese *keiretzu*. Strong linkages in the *keiretzu* are instrumental to the creation of strong networks but prevent the redistribution of agents across different networks. In these networks reconfiguration is centrally controlled, thereby strongly limiting the reshuffling of 'genes'. In a way the flexibility that results from allowing agents to develop local responses to new situations (either the emergence of a new niche or the exploitation of the possibilities open by a new technology) is guaranteed by the fact that weak linkages allow information to flow, new configuration to be tried and consequently new knowledge to be created. New configurations of modules can experiment with prototypes resulting from the quick reshuffling of modules, which, on their side, are not as frequently modified.

The concept of evolvability as presented in this section has to do with a capability of the specific whole to create innovations that are likely to be favourably selected by the internal and external environment. Clusters achieve this by decoupling the level of elementary techniques embedded at the firm's level from the level of final products that instead result from the combination of elementary techniques in transient networks organised *ad hoc*. The concept of NK landscape[590] is useful here. Let's consider any final product produced within the cluster as a phenotypic implementation of a genotypic architecture of elementary techniques (our 'genes'). Genes are owned by individual firms. Firms cluster into networks according to their genes, that is, according to a map of emergent complementarities. Genes represent indivisible (at least at the moment of utilisation) set of processes. Final products can be seen as specific configuration of genes, called chromosomes, in which an ordered sequence of genes specify processes to be carried out and indicate their relative priority. In a NK landscape each chromosome occupies a position on a virtual map, in which the height of the position renders the chromosome's fitness. Adjacent points differ by the value of a single gene. If we put together all possible combinations of genes we end up with a map in which peaks

---

[590] (Kauffman 1995)

indicate points of high fitness and valleys bad solutions. A rational approach to searching for fit solution in the NK landscape is to try to locate the highest peaks. Traditional approaches to new product development involve fixing a basic design and experimenting around the basic variables, changing parameters one at a time to study the effect of the variation on the whole. This is effective in climbing a peak but constraining because peak-to-peak exploration is forbidden. If the landscape is uncorrelated, that is, if the presence of the global peak can not be identified outside a *probe and learn* or *learning-by-doing* approach, moving from peak to peak involves an initial reduction in fitness (going downhill), which would cause the re-orientation of the R&D effort in a new direction. The optimisation procedure works well with a highly correlated landscape, but it is of little efficacy with uncorrelated landscape, or in terms of innovation with radical innovation. From this point of view redundancy, weak linkages and compartmentalisation ensure a different innovation style. Whereas the changes at the level of elementary technique are controlled by single firms and are therefore more likely to be carried out in an incremental fashion following optimisation procedures, in which the search takes place along adjacent peaks (and once the top of the peak is reached it is unlikely to proceed further), changes at the chromosome level instead take place by a recombinant technique at the network level. The reconfiguration of agents in a new network mix the 'genes' with the result of forming a new chromosome, which is likely to be located far away in the NK landscape from the previous one. Suddenly the constraints associated with the optimisation procedure are eliminated. The system can explore areas of the NK landscape that would have been inaccessible to the optimisation procedure. Thanks to recombination and modular 'genes' the system can literally jump from peak to peak. Also, a system based on optimisation at the 'gene' level and recombination at the 'chromosome' level achieves a paradoxical result: it combines an incremental innovation approach at the 'gene' level that ensures a slow rate of change for homeostatic robustness with a radical/architectural innovation style at the level of 'chromosome' that ensures fast change and the exploration of areas negated to incremental innovation. The paradox lies in the fact that the same mechanisms that ensure robustness and incremental change enable fast evolutionary change at the higher level.

The type of innovation, made possible by compartmentalisation, redundancy and weak linkages, can be either intentional or unintended. The latter case is extremely interesting as it opens the dimension of exaptation (this concept has been introduced in chapter 3). Exaptation is the fortuitous discovery of new applications, either at the process or product level, for existing capabilities when exposed to new contexts. The fast metabolism resulting from the acceleration of coevolution in the process of modules reconfiguration creates exactly such conditions. Each new network enables the exposure of existing modules or interconnected modules to new problems/opportunities. The NK interpretation of the role of recombination clarifies an almost mysterious aspect of exaptations and in general of the long-standing problem of the spatial agglomeration of innovations. Exaptations are pseudo-random events in the sense that their occurrence can not be predicted but at the same time their statistical occurrence is linked to the establishment and persistence of some enabling conditions, which, on the one hand drive the system towards a modular structure, and, on the other hand, force the modules to interact with each other in new architectures. Exaptations take place when modularisation creates the conditions for reconfiguration and at the same time drives the cost of experimenting down to a level that the owners of single modules can deal with.

The analogy with the Darwinian evolution should not be pushed too far. Gene reshuffling is purely random in biological species. In clusters however, the process of recombination is a mix of market-push and technology-push. In either case it is not purely random. This implies that the process of change is provided with an emergent direction that makes use of intelligence at the agents and at the network level, therefore ensuring that the rate of success of the mutations generated by the evolutionary incremental/radical innovation mix is higher than the purely random one. This sheds light on evolvability as the capability of generating rapid evolutionary changes with higher probability of success than random processes. In conclusion, a cluster is a specific type of *propagating organisation,* which optimises internal diversity and connectivity to achieve simultaneous homeostasis and rapid evolutionary changes.

7.3 THE CONTRIBUTION TO KNOWLEDGE

Although the concept of diversity often appears in the business, management and economics literature there are very few works that focus directly on diversity to explain the dynamic and structural properties of social systems. I have tried in this work to show that diversity is indeed central by linking diversity to different concepts such as power law distributions, self-organisation, autocatalytic cycles and connectivity. In particular I think that the following points are important.

First, I support Allen's[591] general point of view that self-organising and evolutionary systems require agent's micro-diversity to take place. I show that at least in the case of industrial clusters self-organisation does not take place among clones but on the contrary it requires the presence of micro-diversity among agents. By linking self-organisation and diversity to power law distributions, my research grounds self-organising dynamics into network theories. I show that power law distributions maximise the range of expression of the variable under analysis. Hence, it maximises the diversity of expression of the variable under consideration. More interestingly by comparing structural and dynamic power law variables of the same system, one notices how the emergence of self-organised criticality (dynamic power law) is rooted in a spatial power law distribution of a related variable. Self-organised criticality, rank-size rule and scale-free networks theories seem to be three related aspects pointing to a common underlying pattern. I suggest that the underlying pattern is the *edge of chaos* dynamic, which is the end result of a self-organising dynamic. The correlation between self-organising systems and power law distributions has been proven in the case of the unit of analysis of my research, that is, the Italian industrial agglomerations.

Second, I propose a general model of development of an industrial cluster, based on the mutual interaction between social and economic autocatalytic cycles. Several of the basic blocks that form the cycles are borrowed from the literature on industrial clusters, complexity, innovation and knowledge management. The contribution of my thesis consists in organising them into a general architecture of interactive loops, which I called the economies of diversity. In particular I introduce the distinction between modular innovation at the agent level and architectural (or combinatorial) innovation at the network level and show that the cluster constitutes an appropriate organisational form to manage the tension and dynamics of simultaneous modular and combinatorial innovation. Combinatorial innovation responds to a different logic than modular innovation. I introduce a novel framework, derived from evolutionary biology, called exaptational innovation, that helps explain the dilemma between the extreme geographic concentration of innovations into few hot spots and the apparent randomness (dominated by unintended consequences) of the innovation process in these places.

[591] (Allen 2001)

Third, starting from the latest Kauffman's ideas[592] on the autocatalytic properties of diversity I illustrate how the loops of the economies of diversity are based on the expansion of a specific type of diversity. Taken all together the loops provide an example and a formalisation of Kauffman's *"propagating organisation"*. As the loops generate self-organising systems, it follows that these systems must have a higher diversity than hetero-organised systems. I show however that the relevant concept is not diversity *per se* (as large part of the literature does) but instead what I called systemic diversity, that is, the product of diversity and connectivity. My thesis provides an empirical and original way to measure diversity and connectivity and demonstrates that self-organising systems optimise systemic diversity more than other types of systems.

7.4 SOME IDEAS FOR FURTHER RESEARCH

This thesis opens new directions of research in social systems. Few will contest the statement that economic systems of production and innovation are complex as they are richly interconnected and are underpinned by multiple feedback loops which give rise to unexpected behaviours. However, if the statement is well accepted and documented[593], there are few techniques to discriminate between complex and less complex systems, between systems dominated by self-organisation and systems where hierarchical planning plays a more dominant role.

Drawing heavily from complexity theory and linking it with theories on industrial clusters, I have attempted to sketch a theory of endogenous development of industrial clusters. Clusters develop and grow around a set of autocatalytic loops, defined as economies of diversity. I have also shown how network theories can be used to differentiate between self-organising and hetero-organised systems and how Kauffman's idea of the *propagating organisation* can be formalised in terms of connectivity-diversity product.

The empirical and theoretical work carried out in this thesis opens multiple directions of inquiries. I have divided them into two mains lines, concerned respectively with power laws and innovation.

---

592 (Kauffman 2000)

593 (Krugman 1996)

7.4.1 SYSTEMIC DIVERSITY AND POWER LAWS

Self-organising system clusters exhibit properties that follow a power law distribution. I have demonstrated this point by measuring the size of firms (in terms of number of employees) against their frequency and showing that the resulting distribution is power law. This is however a rather limited case because first, it measures a static property and second, it relies only on one of the three network theories, that is, the rank-size rule. I have provided some theoretical arguments that suggest the equivalence between the three classes of power law phenomena: rank-size rule, scale-free networks and self-organised criticality. The demonstration of the equivalence between the three theories however is not given in this work and is unknown to the author. An interesting expansion of the work started in this thesis is therefore the demonstration of the equivalence of the three approaches to power law in industrial networks. An empirical way to do this is the following. First, select two classes of networks, the former characterised by a mixture of self-organisation and hierarchy skewed toward self-organisation and the latter by a mixture skewed toward hierarchy (or aggregates). Second, isolate an appropriate set of network's variables regarding properties of nodes, links and system's dynamic[594]. Third, plot distribution graphs and test the power law hypothesis in the three cases. If self-organisation correlates positively with a power law in all of the three cases, then this implies that rank-size rule, scale-free networks and self-organised criticality theories are equivalent. This approach would provide an empirical demonstration that self-organising systems are characterised by power laws. Self-organisation brings the systems to the regime of self-organised criticality and is evidenced by a scale-free network distribution in connectivity and by a rank-size power law distribution in the relevant nodal variables.

An alternative line of analysis consists in applying the lessons of the Scheinkman-Woodford model to the connectivity of agents in a supply chain. The already mentioned resilience of industrial clusters to fluctuations in demand may well be a consequence of the power law distribution of relevant variables and the consequent self-organised critical state of the whole system. Why? A system characterised by a distribution dominated by few highly connected hubs and many relatively poorly connected nodes (scale-free

[594] the three types of variables are respectively object of one of the three network theories: rank-size rule, scale-free networks and self-organised criticality

network) is highly resistant to random removal of nodes[595]. In fact, as the hubs are statistically few, the probability of removing one by a random attack is negligible. Therefore the high mortality caused by an industrial crises is less likely to affect the crucial agents through which large part of the connectivity of the system is ensured, that is, the main ganglia of the 'nervous system', which regulates the flows of good and knowledge. The conclusion we may draw from such research is that adaptability is not the only criterion which explains the rate of survival of industrial clusters[596]. Systemic diversity is another factor that preserves industrial clusters' survivability. In fact, connectivity is preserved by the resilience of power law connectivity distribution against random selection, diversity instead by its autocatalytic properties. Until the causes that give rise to the expansion of systemic diversity are not directly under attack (as I have shown in the final section of the story of Meadow in the introduction), then the survivability of a cluster is guaranteed by adaptability and systemic diversity.

Another critical test for the correlation between self-organisation and power law distributions consists in the study of power law distributions and systemic diversity for systems that evolve toward a self-organising network model. This research analyses the properties of industrial agglomerations at one moment in time. If the hypotheses advanced in this work hold true, then the proportionality between closeness to a power law and the degree of self-organisation should be observed in the transition of an industrial agglomeration from an aggregate to a cluster. In parallel it should also be observed an increase in systemic diversity. Moreover, if the assumption is made that it is empirically possible to identify the threshold that divides hetero-organised systems from self-organised systems, then the corresponding change in closeness to a power law and increase in systemic diversity are expected to show a radical non-linear change.

7.4.2 INNOVATION AND SYSTEMIC DIVERSITY

The idea of the propagating organisation is based on the exploration of systemic diversity. Self-organising structures of order survive by exploring the *adjacent possible*. The exploration performs the crucial role of keeping the system far from equilibrium. As

---

[595] (Barabasi 2002)

[596] I support adaptability as a cause in the final part of my conclusions chapter by suggesting that self-organising systems are characterised by a higher evolvability than non-self organising systems. See section 7.2.2

Nicolis and Prigogine write: "n*on equilibrium reveals the potentialities hidden in the nonlinearities, potentialities that remain dormant at or near equilibrium*"[597]. The process of exploration is therefore not an accessory dimension of the cluster but the organisational principle that allows the system to explore the potentialities that are negated to a system under equilibrium and consequently to maintain its structure.

This line of reasoning has deep implications in terms of innovation policy and studies. Technological and organisational innovation becomes the process by which the propagating organisation maintains the out-of-equilibrium connectivity-diversity product in balance between the two runaway mechanism of standardisation (diversity and connectivity-reducing) and diversification (diversity-increasing).

A fascinating line of research could arise around the following research questions:

- Is the capability of innovation (or adaptation) of an economic and social system dependent upon the system's degree of diversity and web of connectivity? Therefore can the study of the conjugated variables diversity-connectivity shed light on the dynamic of innovation?
- What type of innovation is more strongly affected by systemic diversity?
- How can organisations use the dynamics of self-organising networks to achieve the portfolio of innovations they strive for? What types of organisational structural changes are required?

The unit of analysis of the research would be networks of organisations and organisations as networks. These systems present different mixtures of self-organisation and hierarchical properties. In particular an interesting comparative research could analyse:

- Industrial clusters as distributed networks aggregated around localities;
- Virtual networks and community of economic and social agents (for instance open source communities);
- Firms as organisational network

---

[597] (Nicolis and Prigogine 1989) p.60

Appendix

8.1 APPENDIX 1 - MATHLAB PROGRAM: CALCULATION OF GENERIC AND SYSTEMIC DIVERSITY

```
%_________CALCULATION OF DIVERSITY AND SYSTEMIC DIVERSITY________________
%
%
%
%  --------------------------------DIGITISATION------------------------
-
%
%transform a vector of x digit numbers into a matrix where each number
is split into the
%constituent digit

% serves to reduce number of 1 digit

% n declare how many numbers are to be transformed into vectors (rows
of final matrix)
% m number of digits
% the second loop is inverted to get the digits ordered in the vector
%  z(i)contains the individual digits; each iteration of i adds one
more digit
% for each i a is reduced of 1 digit dividing by 10 and subtracting
the decimal part
% the second i loop saves the z values into the H matrix

clear
g=10;         %division factor to reduce digit
m=3;          %number of digit

[SIC,n,Dens]=SICrnd600_biel;

for j=1:n
      for i=m:-1:1
      temp=SIC(j);          %sic: vector containing n industrial codes
      resto=rem(temp,g);
      z(i)=resto;
      SIC(j)=temp/g-resto/g;
      end
   for i=1:m
      H(j,i)=z(i);    %sic stored as single digits
   end
end
H;
%
%
%
%---------------------------DIFFERENCE------------------------------
%
%CALCULATION OF DIFFERENCE BETWEEN ANY TWO ATECO (SIC)

% n number of rows. Each row represents a SIC
% m number of digits
%  z temporary allocation of a row for sake of comparison with all
others. The difference between
```

```
% any one row and the others will constitute a columns of the
final n x n matrix
% a name of matrix with initial data SICs


digit1=2^(m-1); %value of weight
digit2=2^(m-2);
  digit3=2^(m-3);

s=1;                                   %initialisation of column index
for weight nxn matrix
for i=1:n
     x=H(i,:);                         %extraction of row
       for j=1:n                       %summatory over same index for
difference
          diff=x-H(j,:);               %difference; it includes
difference with itself
          round(diff);                 %make sure that round integer are
returned
          if round(diff(1))~=0
          K(j,s)=digit1;               %set difference matrix
          elseif round(diff(2))~=0
          K(j,s)=digit2;
          elseif round(diff(3))~=0
          K(j,s)=digit3;
          end
       end
   s=s+1;
end

distance=sum(sum(K)'); %calculates the sum of differences between
atecos

distance_exp=distance/2 % symmetric matrix along diagonal
%
%
%
%
% ----------------- CALCULATION OF DIVERSITY INDEX ------------------

% diversity index is a scalar. the index is the product of the
difference matrix,
% times frequency of sic i multiplied freq of sic
% Dens represents frequency of firms having a certain sic


for i=1:n
   for j=1:n
      ii=i;                     %check
      jj=j;                     %check
      diff=K(i,j);              %difference matrix: represents the
weight assigned to the difference of any two ateco
      fi=Dens(i);               %frequency of ateco i
      fj=Dens(j);               %frequency of ateco j
      C=diff*fi*fj;             %product matrix
      Divers(i,j)=C;            %values are stored in matrix divers

   end
end

% sum over all elements of matrix. Sum returns a vector containing the
sums of columns.
% taking the transposed matrix of the sum and summing again we get the
final scalar, called somma

diversity=sum(sum(Divers)');
```

```
  diversity_exp=diversity/2                          % symmetric
matrix along diagonal


%
%
%
%
%   _________Extraction Connectivity matrix from IO__________________

% Conn is extracted from IO eliminating rows and columns with SIC
external to specific TtWA

[Conn]=Conn_matrix;
%
%
%
%
%___________Calculation of connectivity index________________________

% conn index represents the sums of all connectivity weight of TtWA

ConnIndex=sum(sum(Conn)') %calculates the sum of differences between
atecos
%
%
%
%
%____Calculation of modularity index (systemic diversity)____________

% mod index represents the product of diversity times connectivity
index

for i=1:n
   for j=1:n
      ii=i;
      jj=j;
      diff=K(i,j);
      c=Conn(i,j);
      fi=Dens(i);
      fj=Dens(j);
      mod=diff*c*fi*fj;
      MOD(i,j)=mod;
   end
end
  ModIndex=sum(sum(MOD)')     %calculates the sum of differences
between                                                      atecos
```

8.2 *APPENDIX 1 - MATHLAB PROGRAM: EXTRACTION OF I-O MATRIX FOR CLUSTER* K

```
% ____ CALCULATION OF SPECIFIC IO TABLE FOR SPECIFIC SLL ____________
%
%
%
% select from general IO matrix a sub-matrix (Conn) with SIC of
specific SLL
%
%
%
function [Conn]=Conn_matrix;
[IO,io]=IO_mt;                % general IO matrix
[SIC,n,Dens]=SICrnd600_biel;  %specific SLL for which specific IO
matrix is extracted
for i=1:io
   for j =1:n
      if IO(i,1)==SIC(j,1)
         C(j,:)=IO(i,:);
      end
   end
end
for i=1:io
   for j =1:n
      if C(1,i)==SIC(j,1)
         Conn(:,j)=C(:,i);
      end
   end
end
Conn(1,:)=zeros(1,n);
  Conn(:,1)=zeros(n,1)
```

References

Abernathy, W. J. and J. M. Utterback (1978). Patterns of industrial innovation. Managing strategic innovation and change: readings in the management of innovation. W. Moore and M. Tushman, Pitman: 141-148.

Adamic, L. A. Zipf, power-laws, and Pareto - a ranking tutorial, Internet Ecologies Areas.

Allen, P. (1997). *Cities and regions as self-organising system*. Amsterdam, Gordon and Breach Science Publisher.

Allen, P. (2001). "A complex systems approach to learning in adaptive networks." International Journal of Innovation Management **5**(2): 149-180.

Almeida, P. and B. Kogut (1997). "The exploration of technological diversity and the geographic localisation of innovation." Small Business Economics **9**: 21-31.

Amin, A. and K. Robins (1990). Industrial districts and regional development: limits and possibilities. Industrial districts and inter-firm co-operation in Italy. F. Pyke, G. Becattini and W. Sengenberger. Geneva, International Institute for Labour Studies: 185-219.

Amin, A. and N. Thrift (1992). "Neo-Marshallian nodes in global networks." International Journal of Urban and Regional Research **16**(4): 571-587.

Anderson, P. and M. L. Tushman (1997). Managing through cycles of technological change. Managing strategic innovation and change. M. L. Tushman and P. Anderson. Oxford, Oxford University Press.

Andriani, P. (2001). "Diversity, Knowledge and Complexity Theory: Some Introductory Issues." International Journal of Innovation Management **5**(2): 257-274.

Arthur, B. "Positive feedback in the economy." Scientific American.

Arthur, M., R. J. DeFillippi, et al. (2001). "Careers, Communities, and Industry Evolution: Links to Complexity Theory." International Journal of Innovation Management- **2**(2): Forthcoming.

Arthur, W. B. (1996). "Increasing returns and the new world business." Harvard Business Review **July-August**.

Arthur, W. B., S. N. Durlauf, et al., Eds. (1997). The economy as an evolving complex system. Reading, MA., Perseus.

Ashby, R. (1956). Introduction to Cybernetics. New York, John Wiley & Sons.

Axelrod, R. (1984). The evolution of cooperation. New York, Basic Books.

Axelrod, R. (1997). The complexity of cooperation: agent-based models of competition and collaboration. Princeton, Princeton University Press.

Aydalot, P. and D. Keeble, Eds. (1988). High technology industry and innovative environments the European experience. London, Routledge.

Bagnasco, A. (2001). Teoria del capitale sociale e "political economy" comparata. Il capitale sociale: istruzioni per l'uso. B. e. Al. Bologna, il Mulino: 77-104.

Bak, P. (1997). How Nature Works : The Science of Self-Organized Criticality, Oxford University Press.

Bak, P. and K. Chen (1991). "Self-organised criticality." Scientific American(January).

Barabasi, A. L. (2002). Linked: the new science of networks. Cambridge, MA, Perseus Publishing.

Becattini, G. (1990). The Marshallian industrial district as a socio-economic notion. Industrial districts and inter-firm co-operation in Italy. G. B. F. Pyke, and W. Sengenberger. Geneva, International Institute for Labour Studies: x, 237 p ill 24 cm.

Becattini, G. (1997). Prato: Storia di una citta' - Il distretto industriale (1943-1993). Firenze, Le Monnier.

Becattini, G. (1998). *Distretti industriali e made in Italy*. Torino, Bollati Boringhieri.

Bellandi, M. (1995). Economie di scala e organizzazione industriale. Milano, Franco Angeli.

Belussi, F. (1999). "Policies for the development of knowledge-intensive local production systems." Cambridge Journal of Economics **23**: 729-747.

Best, M. H. (1990). The new competition institutions of industrial restructuring. Cambridge, Mass, Harvard University Press.

Boari, C., V. Odorici, et al. (2001). Clusters and rivalry: does localisation really matter? Academy of Management.

Boisot, M. (1998). *Knowledge Assets: securing competitive advantage in the information economy*. Oxford, Oxford University Press.

Bronzini, R. (2000). Sistemi produttivi locali e commercio estero: un'analisi territoriale delle esportazioni italiane. Modelli di sviluppo locale: produzione, mercati e finanza: I risultati di una ricerca della Banca d'Italia sui distretti industriali italiani. L. F. Signorini. Roma, Meridiana Libri.

Brown, J. S. (2000). The social life of information. Boston, Harvard Business School Press.

Brown, J. S. and P. Duguid (1991). "Organisational learning and communities-of-practices: toward a unified view of working, learning and innovation." Organisation Science **2**(1): 40-57.

Brusco, S. (1982). "The Emilian model: productive decentralization and social integration." Cambridge Journal of Economics **6**: 167-184.

Brusco, S. and S. Paba (1997). Per una storia dei distretti industriali italiani dal secondo dopoguerra agli anni novanta. Storia del capitalismo Italiano dal dopoguerra ad oggi. F. Barca. Roma, Donzelli.

Buchanan, M. (2000). Ubiquity the science of history - or why the world is simpler than we think. London, Weidenfeld & Nicolson.

Burt, R. (1992). Structural Holes: the social structure of competition. Cambridge MA, Cambridge University Press.

Cairncross, F. (1998). The death of distance : how the communications revolution will change our lives. London, Orion.

Callon, M. (1991). Techno-economic networks and irreversibilities. A sociology of monsters essays on power, technology and domination. J. Law. London, Routledge.

Camagni, R. (1991). Innovation networks: spatial perspectives. London, Belhaven Press.

Campbell, J. (1982). Gramatical Man. New York, Simon and Schuster.

Cannari, L. and L. F. Signorini (2000). Nuovi strumenti per la classificazione dei sistemi locali. Modelli di sviluppo locale: produzione, mercati e finanza: I risultati di una ricerca della Banca d'Italia sui distretti industriali italiani. L. F. Signorini. Roma, Meridiana Libri.

Capra, F. (1988). Uncommon Wisdom. New York, Simon & Schuster.

Capra, F. (1996). The web of life a new synthesis of mind and matter Fritjof Capra. London, HarperCollins.

Castells, M. (2000). The rise of the network society. Oxford, Blackwell.

Castells, M. and P. Hall (1994). Technopoles of the world the making of twenty-first-century industrial complexes. London, Routledge.

Casti, J. L. (1994). Complexification explaining a paradoxical world through the science of surprise. London, Abacus.

Casti, J. L. (1997). Would-be worlds: how simulation is changing the frontiers of science. New York, John Wiley.

Cavalli-Sforza, L. L. (1995). The great human diasporas: the history of diversity and evolution Luigi Luca Cavalli-Sforza and Francesco Cavalli-Sforza translated from the Italian by Sarah Thorne. Reading, Mass., Addison-Wesley.

Chenery, H. and T. Watanabe (1958). "International comparison of the structure of production." Econometrica **26**: 487-521.

Christensen, C. M. (1997). The Innovator's dilemma, Harvard Business School Press.

Christensen, C. M., M. Raynor, et al. (2001). "Skate to where the money will be." Harvard Business Review **November**.

Cilliers, P. (1998). *Complexity and Postmodernism*. London, Routledge.

Cilliers, P. (2000). "Rules and complex systems." Emergence **2**(3): 40-51.

Cipolla, C. M. (1970). An economic history of world population. London, Penguin.

Coase, R. (1937). "The nature of the firm." Economica **4**: 386-405.

Cohen, W. M. and D. A. Levinthal (1990). "Absorptive capacity: a new perspective on learning and innovation." Administrative Science Quarterly **35**(1): 128-153.

Coleman, J. (1990). Foundations of social theory. Cambridge, MA, The Belknap Press of Harvard University Press.

Cringely, R. X. (1996). Accidental empires. London, Penguin.

Cusumano, M. A., Y. Mylonadis, et al. (1997). Strategic manouvring and mass-market dynamics: the triumph of VHS over Beta. Managing strategic innovation and change. M. L. Tushman and P. Anderson, Oxford.

Daniel, W. W. and J. C. Terrel (1989). Business statistics. Boston, Houghton Mifflin Company.

David, P. A. (1985). "Clio and the Economics of QWERTY." American Economic Review **75**(May): 332-7.

Dawkins, R. (1989). The selfish gene. Oxford, OUP.

Dawkins, R. (1996). Climbing mount improbable. London, Viking.

Diamond, J. (1997). Guns, Germs and Steel. The Fates of Human Societies. New York, W.W. Norton and Company.

Dietzenbacher, E. (1992). "The measurement of inter-industry linkages: key sectors in the Netherland." Economic Modelling(October): 419-437.

Dosi, G. (1988). The Nature of the innovative process. Technical Change and Economic Theory. G. D. e. Al. London, Pinter: 221-39.

Dosi, G. and [et al.] (1988). Technical change and economic theory. London, Pinter.

Dosi, G. and L. Orsenigo (1985). Order and change: an exploration of markets, institutions and technology in industrial dynamics, Brighton: SPRU.

Dunning, J. H. (1988). Multinationals, technology and competitiveness. London, Unwin Hyman.

Duranton, G. and D. Puga (1999). Diversity and Specialisation in Cities: Why, where and when does it matter?, CEPR discussion paper.

Duranton, G. and D. Puga (2001). Nursery Cities: Urban diversity, process innovation, and the life-cycle of products, CEPR discussion paper.

Duris, P. and G. Gohau (1999). Storia della biologia. Torino, Einaudi.

Eigen, M. (1971). "Molecular self-organization and the early stages of evolution." Quarterly Review of Biophysics **4**(2&3).

Eldredge, N. (1995). Reinventing Darwin the great evolutionary debate. London, Weidenfeld and Nicolson.

Eldredge, N. and I. Tattersall (1982). The myths of human evolution. New York, Columbia University Press.

Engelstof, S., C. Jensen-Butles, et al. (2002). The economics of industrial clusters and an examination of their performance in Denmark. Regional Science Association International, Brighton.

Ergas, H. (1986). Does technology policy matter? Brussels, Centre for European Policy Study.

Evans, P. and T. Wurster (1999). Blown to bits, Harvard Business School Press.

Evans, P. B. and T. S. Wurster (1997). "Strategy and the new economics of information." Harvard Business Review(Sept-Oct): 71-82.

F. Pyke, G. Becattini, et al., Eds. (1990). Industrial districts and inter-firm co-operation in Italy. Geneva, International Institute for Labour Studies.

Feldman, M. (1999). "The new economics of innovation, spillovers and agglomeration: a review of empirical studies." Economics of Innovation and New Technology **8**: 5-25.

Feynman, R. P., R. B. Leighton, et al. (1963). The Feynman lectures on physics. Reading, Mass, Addison-Wesley.

Fujita, M., A. Venables, et al. (1999). *The spatial economy: cities, regions and international trade*. Cambridge, Mass, MIT Press.

Fukuyama, F. (1995). Trust the social virtues and the creation of prosperity. London, Hamish Hamilton.

Garreau, J. (1992). Edge City. New York, Anchor Books.

Gell-Mann, M. (1994). The quark and the jaguar: adventure in the simple and the complex. New York, W.H. Freeman and Company.

Gerhard, J. C. and M. Kirschner (1997). Cells, embryos and evolution, Blackwell.

Gibbons, M. and J. Metcalfe (1988). Technology, variety and the organisation: a systematic perspective on the competitive process. Research on technology and innovation management policy. Rosembloom and Burgelman. Manchester, PREST, University of Manchester.

Gilbert, G. N. and K. G. Troitzsch (1999). Simulation for the social scientist. Buckingham, Open University Press.

Gladwell, M. (2002). The tipping point: how little things can make a big difference, Abacus.

Glaeser, E., H. Kallal, et al. (1992). "Growth in cities." Journal of Political Economy **100**(6): 1126-1152.

Gleick, J. (1987). Chaos : making a new science. New York, Viking Penguin.

Gleick, J. (2002). What just happened. Michigan, Abacus.

Goodwin, B. (1994). How the leopard changed its spots: the evolution of complexity. London, Orion Books.

Gould, J. (1977). Ever since Darwin: reflections on natural history. New York, Norton.

Gould, J. S. (2000). Wonderful Life: the Burgess Shale and the nature of history. London, Vintage.

Gould, S. J. (1996). Full House: the spread of excellence from Plato to Darwin. New York, Crown Publisher.

Gould, S. J. (1997). The panda's thumb of technology. Managing strategic innovation and change. M. L. Tushmann and P. Anderson. Oxford, Oxford University Press.

Gould, S. J. and N. Eldredge (1977). "Punctuated equilibria: the tempo and mode of evolution reconsidered." Paleobiology **3**: 115-151.

Gould, S. J. and E. S. Vrba (1982). "Exaptation - a missing term in the science of form." Paleobiology **8**: 4-15.

Grabher (2000). "Speed of creativity: heterarchy in the advertising industry." Environment and Planning A (to be published).

Grabher, G. and D. Stark (1997). "Organising diversity: evolutionary theory, network analysis and postsocialism." Regional Studies **31**(5): 533-544.

Granovetter, M. (1985). "Economic Action and Social Structure: the Problem of Embeddedness." American Journal of Sociology **78**: 1360-1380.

Granovetter, M. (1992). Problems of explanations in economic sociology. Networks and organisations: structure, form, and action. N. Nohria and R. G. Eccles. Cambridge, MA, Harvard Business School Press.

Gray, M., E. Golob, et al. (1996). "Big firms, long arms, wide shoulders: the "hub and spoke" districts in the Seattle region." Regional studies **30**(7).

Gundling, E. (2000). The 3M way to innovation: balancing people and profit. Tokyo, Kodansha International.

Hall, R. (1999). "Rearranging Risks and Rewards in a Supply Chain." Journal of General Management **24**(3): 22-32.

Hall, R. and P. Andriani (1999). "Operationalising Knowledge Management Concepts: The Development of a Technique for Sharing Knowledge in New Product Development Projects." International Journal of Innovation Management **3**(3): 307-335.

Hardin, G. (1968). "Tragedy of the commons." Science **162**: 1243-48.

Harrison, B. and M. e. a. Kelley (1996). "Specialization Versus Diversity in Local Economies: The Implications for Innovative Private-Sector Behavior." Cityscape **2**: 61-93.

Henderson, R. and K. Clark (1990). "Architectural innovation: the reconfiguring of existing product technologies and the failure of established firms." Administrative Science Quarterly **35**(1): 9-30.

Henderson, V. (1986). "Efficiency of resource usage and city size." Journal of Urban Economics **19**: 47-70.

Henderson, V., A. Kuncoro, et al. (1995). "Industrial development in cities." Journal of Political Economy **103**(5): 1067-1085.

Henry, N. and S. Pinch (1997). A regional formula for success? The innovative region of Motor Sport Valley, An executive summary of a research project on the British motor sport industry funded by the ESRC.

Hirschman, A. O. and A. Otto (1968). The strategy of economic development. New Haven, Yale University Press.

Holland, J. H. (1995). Hidden order: how adaptation builds complexity. Reading, Massachusetts, Addison Wesley.

ISTAT (1997). I sistemi locali del lavoro. Roma, ISTAT.

ISTAT (2000). Tavole intersettoriali dell'economia Italiana. Roma, ISTAT.

Jacobs (1967). The economies of cities. New York, Random House.

Jeong, H., B. Tombor, et al. (2000). "The large-scale organization of metabolic networks." Nature **407**: 651-654.

Juarrero, A. (1999). Dynamics in action. Cambridge, MA, The MIT Press.

Kale, P., H. Singh, et al. (2000). "Learning and protection of proprietary assets in strategic alliances: building relational capital." Strategic Management Journal **21**(3): 217-238.

Kauffman, S. (1995). At home in the universe. Oxford, Oxford University Press.

Kauffman, S. (2000). Investigations. Oxford, Oxford University Press.

Kelley, M. and S. Helper (1997). "Firm Size and Capabilities, Regional Agglomeration, and the Adoption of New Technology." Economics of Innovation and New Technology **forthcoming**.

Kogut, B. (2000). "The network as knowledge: generative rules and the emergence of structure." Strategic Management Journal **21**(Special Issue on Strategic Networks): 405-425.

Krugman (1996). The self-organising economy. Maiden, MA, Blackwell.

Krugman, P. (1997). How the economy organizes itself in space: a survey of the new economic geography. The economy as an evolving complex system. W. B. Arthur, S. N. Durlauf and D. A. Lane. Reading, MA., Perseus: 583p.
Kuhn, T. S. (1996). The structure of scientific revolutions. Chicago, IL, University of Chicago Press.
Kundera, M. (1984). The unbearable lightness of being. London, Faber and Faber.
Laing, R. D. (1970). Knots. London, Tavistock Publications.
Lane, D., Ed. (2001). Complexity and local interactions: towards a theory of industrial clusters. Heidelberg, Springer Verlag.
Leonard, D. and S. Strauss (1997). "Putting your company's whole brain to work." HBR **July-August**: 111-121.
Leonard-Barton, D. (1992). "Core capabilities and core rigidities: a paradox in managing new product development." Strategic Management Journal. Vol.13 **13**.
Levinson, P. (1997). the soft edge: a natural history and future of the information revolution, Routledge.
Lorenzoni, G. and O. A. Ornati (1988). "Constellations of firms and new ventures." Journal of Business Venturing **3**: 41-57.
Lundvall, B.-A., Ed. (1992). National systems of innovation towards a theory of innovation and interactive learning. London New York, Pinter.
Lynn, G. S., J. G. Morone, et al. (1996). "Marketing and discontinuous innovation: the probe and learn process." California Management Review **38**(3).
Malone, T. W. and R. J. Laubacher (1998). "The dawn of the E-lance economy." Harvard Business Review(Sept-Oct): 145-152.
Margulis, L. (1998). The symbiotic planet : a new look at evolution. London, Weidenfeld & Nicolson.
Margulis, L. and D. Sagan (2002). Acquiring genomes. New York, Basic books.
Markusen, A. (1986). Profit cycles, oligopoly and regional development. Totowa, NJ, Rowman and Littlefield.
Markusen, A. (1996). "Sticky places in slippery space: a typology of industrial networks." Economic Geography **72**(3): 293-313.
Markusen, A. and et al. (1991). The Rise of the gunbelt the military remapping of industrial America. New York, Oxford University Press.
Markusen, A., P. Hall, et al. (1986). High tech America the what, how, where, and why of the sunrise industries. Boston, Allen & Unwin.
Marshall, A. (1916). Principles of Economics: An Introductory Volume. London, MacMillan and Co.
Marshall, A. (1919). Industry and trade a study of industrial technique and business organization, and of their influences on the conditions of various classes and nations. London, Macmillan.
Martin, M. J. C. (1994). Managing innovation and entrepreneurship in technology-based firms. New York, John Wiley.
Martin, R. and P. Sunley (2002). "Deconstructing clusters: chaotic concept or policy panacea." Journal of Economic Geography **3**(1).
Marx, K. (1967). Capital. New York, International Publishers.
Maskell, P. (2001). "Knowledge creation and diffusion in geographic clusters." International Journal of Innovation Management **2**(2): Forthcoming.
Maskell, P. and A. Malmberg (1999). "The competitiveness of forms and regions: 'ubiquitification' and the importance of localised learning." European Urban and Regional Studies **6**.
Maturana, H. R. (1998). The tree of knowledge the biological roots of human understanding. London, Shambhala.

Maynard Smith, J. (1974). Models in ecology. London, Cambridge University Press.

Mayr, E. (2001). What evolution is. New York, Basic Books.

McCann, P. and B. Fingleton (1996). "The regional agglomeration impact of just-in-time input linkages. Evidence from the Scottish electronics industry." Scottish Journal of Political Economy **43**(5): 493-518.

McKelvey, B. (2001). "Energising order-creating networks of distributed intelligence: improving the corporate brain." International Journal of Innovation Management **2**(2): 181-212.

McKelvey, B. (2002). Managing coevolutionary dynamics. 18th EGOS Conference, Barcelona, Spain.

McKelvey, B. (2003). Microstrategy from macroleadership: distributed intelligence via new science. Mobilizing the self-renewing organization (forthcoming). A. Y. Lewin and H. Volberda. Thousand Oaks, Ca, Sage.

Midmore, P., M. Munday, et al. (2002). Assessing industry linkages in regional input-output tables. Regional Science Association International, Brighton.

Mokyr, J. (1990). The lever of riches. Technological creativity and economic progress. New York, Oxford University Press.

Mokyr, J. (1991). "Evolutionary biology, technological change and economic history." Bulletin of Economic Research **43**(2): 127 149.

Mokyr, J. (1998). Neither chance nor necessity: evolutionary models and economic history, Princeton University Press.

Monod, J. (1972). Le hasard et la necessite. London, Collins.

Naeem, S. (2002). "Biodiversity equals instability." Nature **416**: 23-24.

Nelson, R. R., Ed. (1993). National innovation systems a comparative analysis. New York Oxford, Oxford University Press.

Nelson, R. R. and S. G. Winter (1982). An evolutionary theory of economic change. Cambridge, MA, Harvard University Press.

Nicolis, G. and I. Prigogine (1989). Exploring complexity: an introduction. New York, W.H. Freeman and Company.

Nishiguchi, T. and A. Beaudet (1998). "Fractal design: self-organised links in supply-chain management." Sloan Management Review **40**(1): 49-59.

Nohria, N. and R. G. Eccles (1992). Networks and organisations: structure, form, and action. Cambridge, MA, Harvard Business School Press.

Ohlin, B. (1933). Interregional and international trade. Cambridge, Harvard University Press.

Pain, S. (2000). "Outrageous fortune." New Scientist **19 February**: 48-49.

Pavitt, K. and P. Patel (1991). "Large firms in production of the world's technology: an important case of non globalisation." Journal of International Business Studies **First Quarter**: 1-21.

Perrow, C. (1992). Small firms networks. Networks and organisations: structure, form, and action. N. Nohria and R. G. Eccles. Cambridge, MA, Harvard Business School Press.

Peters, E. O. (1999). Complexity, Risk, and Financial Markets. New York, John Wiley.

Piore, M. J. (1980). The technological foundations of dualism and discontinuity. Dualism and discontinuity in industrial societies. S. B. a. M. J. Piore. Cambridge, Cambridge University Press.

Piore, M. J. and C. F. Sabel (1984). The second industrial divide: possibilities for prosperity, BasicBooks.

Pool, R. (1997). Beyond engineering : how society shapes technology. New York, Oxford University Press.

Popper, K. R. (1974). Conjectures and refutations : the growth of scientific knowledge. London, Routledge.
Porac, J. and J. A. Rosa (1996). Rivalry, industry models, and the cognitive embeddedness of the comparable firm. Advances in Strategic Management:. A. H. P. Shrivastava, & J. Dutton. Greenwich, JAI Press: 363-388.
Porter, M. E. (1990). The competitive advantage of nations. New York, Free Press.
Porter, M. E. (1991). Michael E. Porter on competition and strategy. Boston, Mass., Harvard Business School.
Porter, M. E. (1998). "Clusters and the new economics of competition." Harvard Business Review(Nov-Dec).
Powell, W. W. (1990). "Neither market nor hierarchy: the network form of organization." Research In Organizational Behaviour **12**: 295-336.
Prahalad, G. K. and G. Hamel (1990). "The core competence of the corporation." Harvard Business Review **68**(3): 79-91.
Pred, A. (1966). The spatial dynamics of the US urban-industrial growth. Cambridge, MA, MIT Press.
Prigogine, I. (1980). From being to becoming : time and complexity in the physical sciences. San Francisco, W.H. Freeman.
Prigogine, I. and I. Stengers (1997). The End of Certainty : Time, Chaos, and the New Laws of Nature. New York, Free Press.
Prigogine, I. I. (1984). Order out of chaos man's new dialogue with nature Ilya Prigogine and Isabelle Stengers foreword by Alvin Toffler. London, Heinemann.
Putnam, R. (1993). Making democracy work. Princeton, NJ, Princeton University Press.
Rasmussen, P. (1958). Studies in intersectoral relations. Amsterdam, North-Holland.
Raup, D. M. (1993). Extinction bad genes or bad luck? Oxford, Oxford University Press.
Raup, D. M. (1999). The nemesis affair a story of the death of dinosaurs and the ways of science. New York, Norton.
Réka, A., H. Jeong, et al. (1999). "Diameter of the World Wide Web." Nature **401**: 130-131.
Rescher, N. (1979). Cognitive Systematization: a system-theoretic approach to a coherent theory of knowledge. Totowa, New York, Rowan and Littlefield.
Resnick, M. (1997). Turtles, termites and traffic jams. Cambridge, MA, The MIT Press.
Riolo, R. L., M. D. Cohen, et al. (2001). "Evolution of cooperation without reciprocity." Nature **414**: 441-443.
Roger, E. M. (1995). Diffusion of innovation. New York, Free Press.
Romer, P. M. (1986). "Increasing returns and long-run growth." Journal of Political Economy **94**(5): 1002-1037.
Rosenberg, N. (1982). Inside the Black Box: Technology and Economics. Cambridge, Cambridge University Press.
Rosenberg, N. (1994). Exploring the black box : technology, economics, and history. Cambridge ; New York, Cambridge University Press.
Rullani, E., Ed. (2001). The industrial cluster as a complex adaptive system. Heidelberg, Springer Verlag.
Sabel, C. F. (2001). Diversity, Not Specialization: The Ties that Bind the (New) Industrial District. Complexity and Industrial Clusters: Dynamics and Models in Theory and Practice. A Conference Organized by
Fondazione Montedison Under the Aegis of Accademia Nazionale dei Lincei, Milan, June 19-20th, 2001.

Sabel, C. F. and J. Zeitlin, Eds. (1997). World of possibilities flexibility and mass production in western industrialization. Cambridge, Cambridge University Press.

Sanchez, R. and R. P. Collins (2001). "Competing - and Learning - in Modular Markets." Long Range Planning **34**: 645-667.

Saviotti, P. P. (1996). Technological evolution, variety and the economy. Cheltenham, Edward Elgar.

Saviotti, P. P. and G. S. Mani (1995). "Competition, variety and technological evolution: a replicator dynamics model." Journal of Evolutionary Economics **5**: 369-392.

Saxenian, A. (1994). Regional advantage: culture and competition is Silicon Valley and Route 128. Harvard MA, Cambridge University Press.

Scheinkman, J. and M. Woodford (1994). "Self-organized criticality and economic fluctuations." The American Economic Review **84**(2): 417-421.

Schelling, T. C. (1978). Micromotives and macrobehavior. New York London, Norton.

Schumpeter, J. A. (1939). Business cycles : a theoretical, historical, and statistical analysis of the capitalist process. New York, McGraw-Hill.

Scott, A. J. (1988). Metropolis: from the division of labor to urban form. Berkeley, University of California Press.

Scott, J. (1991). Social Network Analysis. London, Sage.

Sforzi, F. (1990). The quantitative importance of Marshallian industrial districts in the Italian economy. Industrial districts and inter-firm co-operation in Italy. G. B. F. Pyke, and W. Sengenberger. Geneva, International Institute for Labour Studies: x, 237 p ill 24 cm.

Shapiro, C. (1999). Information rules a strategic guide to the network economy by Carl Shapiro and Hal R. Varian. Boston, MA.,, Harvard Business School Press.

Sigmund, K., E. Fehr, et al. (2002). "The economics of fair play." Scientific American(January): 80-85.

Snow, C. P. (1959). The two cultures and the scientific revolution. Cambridge, Cambridge Univ. P.

Squire, L. R. and E. R. Kandel (1999). Memory : from mind to molecules. New York, Scientific American Library.

Sraffa, P. (1960). Production of commodities by means of commodities : prelude to a critique of economic theory. Cambridge, Cambridge Univ. P.

Stirling, A. (1998). On the economics and analysis of diversity [electronic paper]. Brighton, SPRU.

Storper, M. (1982). "The transition to flexible specialisation in the US film industry: external economies, the division of labor, and the crossing of industrial divides." Cambridge Journal of Economics **13**: 273-305.

Storper, M. (1992). "Regional world of production: learning and innovation in the technology districts of France, Italy and USA." Regional Studies **27**(5): 433-455.

Storper, M. (1997). The Regional World: territorial development in a global economy. New York, Guildford Press.

Tattersall, I. (2001). "How we came to be human." Scientific American(December).

Teece, J. D. (1987). Profiting from technological innovation: implications for integration, collaboration, licensing and public policy. The competitive challenge. D.J.Teece. New York, HarperCollins/Ballinger.

Thomas, A. T. and J. E. Robin (1996). "Making difference matter: a new paradigm for managing diversity." Harvard Business Review(September-October): 79-90.

Trends Business Research Consortium (2001). Business in the UK - a first assessment, Department of Trade and Industry.

Waldrop, M. M. (1992). Complexity, the emerging science at the edge of order and chaos. New York, Simon & Schuster.
Weick, K. E. (1995). Sensemaking in organizations. Thousand Oaks, Sage.
Wenger, E. C. and W. M. Snyder (2000). "Communities of practice: the organizational frontier." Harvard Business Review(January-February): 139-145.
Whitford, J. (2001). "The decline of a model? Changes and responses in the Italian industrial districts." Economy and Society **30**: 38-65.
Wiener, N. (1948). Cybernetics. Cambridge, MA, MIT Press.
Williamson, O. E. (1985). The economic institutions of capitalism firms, markets, relational contracting. New York London, Free Press Collier Macmillan.
Wilson, E. O. (1992). The diversity of life, Belknap Press.